RELIABILITY MODELING OF ULTRA-THIN GATE OXIDE AND HIGH-K DIELECTRICS FOR NANO-SCALE CMOS DEVICES

LOH WEI YIP

B. Eng (Hons), NUS

A THESIS SUBMITTED FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF ELECTRICAL
AND COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2004

Acknowledgments

First and foremost, my deepest gratitude to my supervisors, Associate Professor Cho Byung Jin and Professor Li Ming Fu, who have given me guidance throughout my study in NUS. In particular, it is Assoc. Prof. Cho who have aspired me to reach for the highest standard in my researches and who have tirelessly reviewed and guided me in all my publications. It is with his help that I am able to produce credible results in the area of oxide and high-*K* reliability. Gratitude also goes to Prof. Li, who because of his insight and theoretical expertise is able to guide me to seek for a more theoretical understanding in all my researches. Without Assoc. Prof. Cho and Prof. Li's kind and patient guidance, it would be difficult for me to have completed this thesis.

The advices and guidance of other teaching staffs are also gratefully acknowledged. In particular, Prof. Kwong DL, Assoc. Prof. Yoo WJ, Dr. Zhu CX, Dr Lee SJ, Dr. Yeo YC and Mr. Joo MS, have all given me tremendous help, advices and encouragements. I also wish to express my sincere gratitude to my fellow students in Silicon Nano Device Lab. (SNDL) and Center for Integrated Circuits Failure Analysis and Reliability (CICFAR) who have make my stay in NUS a joyful and meaningful experience. In particular, it is most gratifying to have the support and friendship of Mr. Kim Sun Jung who has so willing lend a helpful hand in all my experiments and Dr. Lim Peng Soon who has joined me in many fruitful discussions in both work and social matters. Thanks also go to all my friends including Wu Nan, Ren Chi, Tony Low, Chee Keong, Ng TH, Zerlinda, Tan KM, Chen JH, Wang YQ, Yu HY, Whang SJ, Park CS, and many others who have went out of their way to teach and assist me in this thesis.

The support and assistance from all the staffs of SNDL and NUS is also gratefully acknowledged. In particular, Mr. Patrick Tang, Mr. Yong Yu Foo, Mr. Goh Thiam Pheng, Mrs. Ho Chiow Mooi, and Mr. Walter Lim have greatly assisted me in all manners of my administrative duties and experiments. Their kind assistances are greatly appreciated.

Last but not least, I wish to dedicate this thesis to my parents, sister, Sook Fen and Gabriel and my dearest Tze Chieg. Without their emotional support, care and concerns, and continuous support and love, I would not have the privilege to even embark on this journey of my life.

Summary

As complementary metal-oxide semiconductor (CMOS) technology advances, the dimensions of its various key device components are scaled downward, from its present day micrometer range and eventually, to its ultimate limit - the nanometer regime. In this aspect, silicon dioxide (SiO_2), which forms the gate insulator for the transistor, is progressively reduced from thick to thin oxide (< 20 Å), ultra-thin (<15 Å) and eventually to high-*K* dielectrics. For high performance logic applications, gate oxide thickness scaling is driven by the need for higher switching speed, transistor drive current and minimization of short channel effects. However, as gate oxide scales to sub-5 nm regime, various reliability phenomena have become increasing prevalent and critical. Quasi-breakdown (QB), which is prevalent in sub-5 nm gate oxides, has become an increasing concern due to its significant impact at low gate voltage and signal noise increases. In the ultra-thin (< 15 Å) regime, gate oxide does not exhibit discrete occurrences of gate leakage current but shows progressive breakdown instead. Moreover, as gate oxide scales even further till sub-nanometer regime, there are increasing evidences that this scaling will be limited by gate dielectric leakage and reliability. At one nanometer, conventional silicon dioxide leakage current at operating voltage, is in the regime of 10 ~ 100 A/cm^2 which may be too high for low power application. As a result, high-*K* dielectrics will be needed by the year 2007 for 65 nm technology node. The breakdown mechanisms, pertaining to different thicknesses of gate dielectrics, have to be well characterized and understood. In the long term, reliability studies for high-*K* will be even more challenging due to its differences in material and electrical properties compared to conventional SiO_2.

For thin gate oxide, in the 30 ~ 45 Å thickness regime, the formation, conduction, and evolution of quasi-breakdown are investigated. Using carrier separation measurements, the electron and hole components of the gate leakage current at onset of QB, are measured and analyzed. Subsequently, bias and thermal annealing are performed on post-QB oxides and disparate responses are observed. By carefully analyzing all the experimental evidences, a unifying defect-induced

breakdown model is presented and verified. As gate oxide scales into ultra-thin regime (< 15 Å), QB becomes masked by the high gate leakage current and subsequent multiple QB spots can be observed and tolerated. A statistical study is conducted on ultra-thin silicon dioxides and a physical model based on multiple quasi-breakdowns is proposed and experimentally verified.

Eventually, high-*K* dielectrics are required for continual gate dielectric scaling. The reliability for high-*K* stacks is examined and a novel technique for stack reliability is presented. Polarity dependent charge trapping in $HfAl_xO_y$ (Hafnium-Aluminum-Oxide) stack is observed and this is correlated to preferential breakdown in the high-*K* and interfacial layer (IL) stack. Using carrier separation measurements, breakdowns in high-*K* stack are attributed to an interface-initiated or bulk layer breakdown in the high-*K* dielectric stacks.

Contents

Acknowledgements i

Summary ii

Table of Contents iv

List of Symbols viii

List of Figures x

List of Tables xx

1. Introduction

1.1 Dimension Scaling and Future Trends of Microelectronics 1

1.2 Summary and Motivation of Thesis 5

1.3 Thesis Outline and Original Contributions 6

References 9

2. Literature Review: Gate Dielectric Degradation

2.1 Impact of device scaling on gate dielectric degradation 11

2.2 Electrical Stress-Induced Degradation and Breakdown 14

2.3 Quasi-breakdown Mechanism 16

2.3.1 Direct Tunneling Model [2.10][2.26]-[2.30] 19

2.3.2 Percolation Model [2.40]-[2.43] 21

2.4 Device Scaling and Dielectric Performance 23

2.5 Ultra-thin oxide Reliability 24

2.6 High-K Dielectrics Reliability 27

2.6.1 High-K charge trapping 28

2.6.2 Stack Reliability 28

2.7 Summary 30

References 32

3. Measurement Setup and Techniques

3.1 Measurement Techniques 41
3.1.1 Direct Current Current-Voltage (DCIV) Technique 41
3.1.2 Charge Pumping (CP) Measurement Technique [3.14] . 44
3.1.3 Comparison between DCIV and CP Technique 46
3.1.4 Carrier Separation 48
3.2 Summary 50
References 51

4. Bipolar Stressing, Trap Generations and QB Mechanism Study

4.1 Introduction 53
4.1.1 Bipolar and Unipolar Current Stressing 55
4.2 Devices and Experimental Setup 56
4.3 Complete Evolution of Oxide Degradation Stages 57
4.3.1 Trap generation in thin gate oxides 59
4.4 Trap Generation and Fluence dependency 59
4.5 Critical level of trap density at onset of QB 62
4.6 Field and Area Dependency at QB 64
4.7 Comparison of QQB for bipolar and unipolar current stressing. .. 66
4.8 Carrier Separation Results 69
4.9 F-N and Direct Tunneling Modeling Equations 73
4.9.1Electron Leakage Current 73
4.9.2Hole dominance leakage current after QB 76
4.10 Proposed Model for QB Mechanism 77
4.11 Summary 81
References 82

5. Effect of Bias and Thermal Annealing on QB and its Mechanism Study

5.1 Introduction 85
5.2 Device and Experimental Setup 86
5.3 Characteristics of Electrical Recovery under Bias and Thermal Annealing 87
5.3.1 Bias Annealing of post-QB oxides 88

5.3.2 Reverse Bias Annealing after QB 90
5.3.3 Thermal Annealing after QB 94
5.3.4 Recoverable and Unrecoverable QB states 99
5.3.5 Combined Annealing Results: Bias and Thermal Anneal 101
5.4 Discussions for bias and thermal annealing 104
5.4.1 Mechanism of thermal and bias anneal 105
5.5 Summary 109
References 110

6. Degradation and Breakdown Mechanism in Ultra-thin Oxides
6.1 Introduction 112
6.2 Device Fabrication and Experimental Setup 115
6.3 Thin Oxide (20 Å) QB Degradation Characteristics 116
6.3.1 Comparison between Thick and Thin oxide: Interface Trap Enhanced Tunneling (ITET) 116
6.3.2 Direct Correlation between interfacial traps and gate leakage 121
6.3.3 Distinction between ITT and ITET 124
6.3.4 Area dependency of ITET 126
6.4 Ultra-thin (<14 Å) oxide Degradation Mechanism 128
6.4.1 Area Dependency of Progressive Breakdown in Ultra-thin Oxides 134
6.5 Modeling of Gate Leakage Current in Ultra-thin Oxides 138
6.5.1 Empirical Experimental Fitting 139
6.5.2 Modeling of Multiples Breakdown Spots 139
6.6 Summary 143
References 144

7. High-K Dielectrics Reliability: Charge Trapping and Breakdown characteristics
7.1 Introduction 148
7.2 Device Fabrication 150
7.3 Carrier Separation and Leakage Path Mechanism 150

7.4 Polarity Dependent Charge Trapping 153
7.5 Experimental Results 154
7.5.1 Charge Trapping in High-K stacks 154
7.5.2 Polarity dependent breakdown characteristics in MOS capacitors 158
7.5.3 Negative CVS : p-MOSFET 161
7.5.4 Positive CVS : n-MOSFET 166
7.5.5 Statistical Breakdown Results 168
7.6 Proposed Charge Induced Breakdown Model 173
7.7 Summary 175
References 176

8. Conclusion and Recommendations
8.1 Conclusions 181
8.2 Recommendations for Future Work 183
References 186

A List of Publications 187

List of Symbols

A_T	Total gate area
c_e^t	Electron capture coefficient
C_{ox}	Oxide Capacitance per unit area
D_{IT}	Interface-state density
E_c	Silicon conduction band edge
E_F	Fermi level
E_g	Silicon bandgap
E_i	Silicon intrinsic energy level
E_{ox}	Oxide electric field
$E_{S,\,max}$	Surface electric field at maximum base recombination current
E_v	Silicon valence band edge
F(t)	Cumulative probability function
g_m	Transconductance
I_B	Base recombination current
$I_{B,max}$	Maximum base recombination current
I_{off}	Off-state transistor leakage current
I_s	Source current
$I_{s/d}$	Current collected from source/drain terminals
I_w	Substrate well current
J_{DT}	Direct tunneling gate current density
J_{FN}	Fowler-Nordheim gate current density
J_g	Gate leakage current density
k	Boltzmann constant
K	Kelvin
K	Dielectric permittivity constant
L	Channel length

m_e	Rest mass of electron
n_i	Intrinsic density of state
N_{IT}	Interface-state density
q	Electronic charge of electron
Q	Charge fluence
QB	Quasi-breakdown
Q_{BD}	Charge-to-breakdown
Q_f	Oxide fixed charge density
Q_{inj}	Injected electron fluence
Q_{IT}	Interface trap charge density
Q_{OT}	Oxide trapped charge density
Q_{QB}	Charge-to-quasibreakdown
R_J	Ratio of current density of degraded spot over fresh current density
T	Temperature (in Kelvin)
$T_{100\% Ig}$	Time-to-100% increment in gate leakage current
T_{BD}	Time-to-breakdown
T_{ox}	Gate oxide thickness
T_{QB}	Time-to-Quasi-breakdown
V_{BE}	Forward voltage bias for base-emitter
V_d	Drain voltage
V_{dd}	Supply voltage
V_{FB}	Flatband voltage
V_g, V_G	Gate voltage
V_{GB}	Gate voltage corresponding to maximum base recombination current
V_{ox}	Voltage drop across oxide/dielectric layer
V_s	Source voltage
V_{th}	Threshold voltage
W	Channel Width
γ	Quantum yield factor
ε_r	Relative permittivity to air
$\bar{x}$	Charge Centroid

List of Figures

Fig. 1.1 Geometry scaling for MOSFET device channel length and equivalent oxide thickness for low power application MOSFETs. Equivalent oxide thickness (EOT) is used instead of physical oxide thickness due to the potential change in dielectrics to high-K material for 65 nm technology node. 2

Fig. 1.2 Trend in long term reliability requirement for MOSFETs. 1 FITs = 1 failure per 10^9 device hours. 4

Fig. 2.1 Hot carrier generation and various current components in n-MOSFET. (After [2.1]) ... 11

Fig. 2.2 Schematics of three possible conduction mechanisms leading to SILC leakage current. (a) Trap-assisted tunneling (TAT) at distinct defect locations with energy relaxation (b) Trap-assisted tunneling at same defect distribution (c) Tunneling and recombination at oxide defect sites (RTAT). (After [2.8]) 12

Fig. 2.3 Illustration of differences between (a) Fowler-Nordheim (FN) and (b) direct tunneling (DT). FN tunneling occurs when $V_{ox} > \Phi_b$ while DT occurs when $V_{ox} < \Phi_b$. 13

Fig. 2.4 Low-voltage conduction mechanism for thin oxide of various oxide thicknesses. (After [2.11]) 14

Fig. 2.5 Anode Hole Injection model with an incident electrons arriving at the anode and transferring its energy to a deep level valence electron and in the process, creating a hole which is then injected back into the oxide. (After [2.15]) 15

Fig. 2.6 Evolution of gate voltage under constant current stress till complete breakdown. At quasi-breakdown (QB), gate voltage drops due to enhanced leakage path with gate voltage magnitude still significantly higher than at complete breakdown. 17

Fig. 2.7 I-V characteristics of gate oxide at various stages of stressing – Fresh, SILC, quasi-breakdown (QB) and complete breakdown. 17

Fig. 2.8 Schematic drawing for (a) electron transport in the ultra thin gate oxide under high field stress (b) Current path in the oxide after quasi-breakdown. (After [2.10]) 20

Fig. 2.9 Schematic illustration of percolation model for intrinsic oxide breakdown based on electrons trap generation. Conduction path is indicated by the shaded spheres. (After [2.33]) 21

Fig. 2.10 Current-voltage characteristics of a MOS capacitor with a 4.2 nm gate oxide. Solid line is fit obtained with a percolation model. (After [2.43]) 22

Fig. 2.11 Simulated and measured Weibull slope β for charge-to-breakdown QBD as a function of oxide thickness. (After [2.40]) 23

Fig. 2.12 Gate leakage current for different oxide thickness at fresh unstressed state and after post-QB. It can be observed that post-QB leakage current for 35 - 45 Å oxides are actually lower than that of the direct tunneling leakage current of 13 - 20 Å at its initial unstressed state. 25

Fig. 2.13 Time evolution of gate current before and after onset of QB for ultra-thin oxides (13.5 Å) under various gate bias CVS. 26

Fig. 3.1 p-MOSFET in top emitter-base configuration with spatial distribution of interface traps and recombination traps centers as shown. 42

Fig. 3.2 Basic experimental setup for DCIV measurement using p-MOSFET. A vertical parasitic p/n/p-BJT is used with p+ source as emitter, n-well as base, and p-substrate as collector. Forward bias condition of V_{EB} = + 0.3 V and V_{BC} = 0 V is applied and base recombination current I_B is monitored as a function of gate voltage V_g sweep. Drain can be connected together with source as shown or floated. 44

Fig. 3.3 Basic experimental setup for conventional charge-pumping measurements on p-MOSFETs. Source and drain are shorted and lightly reversed biased. An ac signal is applied to gate to alternately drive electrons and holes into the interface traps located at the SiO_2-Si interface. During channel accumulation, recombination of free carrier with the trapped charges cause a net DC substrate current – charge pumping current I_{cp} which is proportional to the areal interface trap density.....45

Fig. 3.4 Schematic illustration of CP technique applied to p-channel MOSFETs. (a) Fixed top level V_{gh} with variable V_{gb} and (b) fixed base level V_{gb} and variable V_{gh}. Both configuration shown with the associated charge pumping current versus the variable gate voltage as shown in the lower figures. 45

Fig. 3.5 Basic experimental setup for carrier separation measurement on p-channel MOSFETs under inversion mode. Drain is floated while source and n-well are grounded. Gate voltage is swept from 0 V to negative 3 V. 48

Fig. 3.6 Schematic band diagram (a) of p-channel MOSFET in inversion mode and (b) n-channel MOSFET measured in inversion mode. [After [3.1]].....49

Fig. 4.1 Measured gate voltage due to application of constant bipolar current pulse of current density J = +/- 50 mA/cm^2. 57

Fig. 4.2 Complete evolution of oxide degradation for thin oxide (T_{OX} = 45 Å) under bipolar constant current stress (J_{stress} = +/- 10 mA/cm^2). It can be observed that within QB there are 2 stages – recoverable and unrecoverable QB. (n-MOSFET, W/L = 10/0.2 μm).....58

Fig. 4.3 I-V characteristics of oxides at various stages of stressing conditions – fresh, recoverable QB, unrecoverable QB and complete breakdown. (n-MOSFET, W/L = 10/0.2 μm).....58

Fig. 4.4 Peak recombination current $I_{B,max}$ versus charge fluence for stressing till quasi-breakdown. Since peak $I_{B,max}$ is proportional to interface trap, it can be observed that interface trap density is proportional to charge fluence $Q^{0.52}$ independent of channel area and stressing current density. 60

Fig. 4.5 Voltage shift of peak $I_{B,max}$, V_{GB} versus charge fluence. Since oxide trap density is proportional to V_{GB}, it can be observed that oxide trap density can be divided into 2 region. At charge fluence < 100 C/cm^2, oxide trap is proportional to $Q^{0.043}$ while beyond, oxide trap is proportional to $Q^{0.31}$.....61

Fig. 4.6 Trap generation under constant voltage stressing. Oxide trap N_{OT} is proportional to $Q^{0.18}$ while interface trap N_{IT} is proportional to $Q^{0.27}$. (p-MOSFET, W/L=10/1 μm, V_{stress} = 6.8 V) 62

Fig. 4.7 Critical level of oxide trap at onset of quasi-breakdown for different channel area as shown. Oxides are stressed till QB using CVS at different gate bias. It can be observed that bulk traps as reflected by ΔV_{GB} at onset of QB increase with gate bias, without a single constant level expected for critical bulk defects for percolation model. 63

Fig. 4.8 Critical level of interface traps at onset of quasi-breakdown for different channel area. Oxides are stressed to QB using CVS at different gate bias. It can be observed that a constant level of interfacial traps is obtained irrespective of stressing gate bias. Similar results have already been reported in [4.9].....64

Fig. 4.9 Voltage shift of peak I_B versus stress fluence for different stressing current density (a) J_{stress} = 0.5 mA/cm^2 (b) J_{stress} = 2 mA/cm^2 (c) J_{stress} = 20 mA/cm^2. It can be observed that 2 stages exist for bulk trap generation under CCS. The crossover point is highly dependent on the magnitude of the constant current stress. In the initial stage, bulk trap generation is much slower and highly dependent on gate bias. In the second stage, trap generation is much faster and total traps generated as reflected by the magnitude of voltage shifts appear to be independent of gate bias.....65

Fig. 4.10 Charge to quasi-breakdown for small channel area (< 3 μm^2) using various mode of constant current stress with different polarity injection, including unipolar and bipolar CCS. It can be observed that bipolar stressing results in much lower QQB for the same fluence as compared to unipolar stressing. Each point (differentiated by sample area) is obtained from 5-10 samples with Q_{QB} (63% values) corresponding to the zero level in the Weibull distribution.....67

Fig. 4.11 Charge to quasi-breakdown for big channel area (~100 μm^2) using various mode of constant current injection. For large area samples, gate injection unipolar stress

results in significantly lower Q_{QB} compared to substrate injection. Bipolar stressing for large samples also has low Q_{QB} and appears to be limited by the low Q_{QB} for gate injection under negative gate bias. Results are similar to [4.13]. Each point is obtained from 5-10 samples with Q_{QB} (63% values) corresponding to the zero level in the Weibull distribution. 68

Fig. 4.12 Carrier separation measurement showing gate, source and substrate current component at fresh, unstressed state. (p-MSOFET, T_{ox} = 45 Å, W/L = 10/0.5 μm) 70

Fig. 4.13 Carrier separation measurement at onset of QB, which is attained after 222C/cm^2 of electron fluency. Sample is still in recoverable QB stage and is the same one used in Fig. 4.12. 71

Fig. 4.14 Carrier separation measurement after post-QB stress (additional electron fluency of 38 C/cm^2) within recoverable QB stage. Sample used is the same as Fig. 4.12 71

Fig. 4.15 Carrier separation measurement after post-QB stress but stressed to unrecoverable QB (p-MOSFET, T_{ox} = 45 Å, W/L = 10/1 μm, J_{stress} = 10 mA/cm^2) 72

Fig. 4.16 Experimental quantum yield as a function of gate voltage and resulting electron energy for unstressed gate oxides and post-QB oxides as shown inset. 73

Fig. 4.17 Carrier separation for well current component at fresh and onset of QB state. F-N current is simulated using (3.5) with varying energy barrier, oxide thickness and electric field shift. Good fit observed for experimental I_{well} and F-N current using electric field shift V_{shift} = 2.2 V. 75

Fig. 4.18 Evolution of well current component for post-QB stage under continual stressing. Good fit observed for well current component with simulated F-N tunneling current. (T_{ox} = 45Å , W/L = 10/0.7 μm, p-MOSFET) 76

Fig. 4.19 Evolution of source current component after QB in hole dominant regime. Relatively good fit observed between experimental data at QB and direct tunneling current modeling using (3.6). (T_{ox} = 45Å, W/L = 10/0.7μm, p-MOSFET) 77

Fig. 4.20 A schematic drawing of energy band diagram for localized trap region (LTR) model. Hole trapping at anode results in distortion of energy band and formation of localized trap region (LTR) causing bandgap narrowing near the anode. (a) In initial stage, electron conduction by F-N tunneling. (b) Further stressing extends the LTR, resulting in hole direct tunneling. Electrons and holes are indicted by solid and open circles, respectively. 79

Fig. 4.21 Schematic illustration of evolution of the localized trap region (LTR) formed by deep level trapped holes at various stages of QB (A) at onset of QB, within recoverable QB, LTR is mainly localized at the anode and conduction proceed by direct tunneling of both holes and electrons. (B) at unrecoverable QB, LTR has extended the whole oxide forming a direct conduction path. 81

Fig. 5.1 Schematic illustration of carrier separation measurement setup for electrical bias annealing. 87

Fig. 5.2 Evolution of gate voltage under constant current stress till QB and post –QB positive bias annealing. (p-MSOFET, T_{ox} = 45 Å, W/L = 10/0.5 μm, J_{stress} = 50 mA/cm^2) 88

Fig. 5.3 DCIV spectra of p-MOSFET for Fig. 5.2, with stressing till QB and post-QB. After onset of QB, it can be observed that the recombination current, I_B spectra overlaps with subsequent decrease in the peak amplitude of I_B accompanied by a slight shift of VGB for peak IB to the right. (p-MSOFET, T_{ox} = 45 Å, W/L = 10/0.5 μm, J_{stress} = 50 mA/cm^2) 89

Fig. 5.4 Quantitative DCIV spectra measurement showing IB, max and ΔVGB versus injected fluencies. Bulk and interface trap after QB show no further increment. (p-MOSFET, T_{ox} = 45 Å, W/L = 10/0.7 μm, J_{stress} = 50 mA/cm^2) 89

Fig. 5.5 Evolution of gate voltage under constant current stressing. At QB, reverse bias current J_{rev} = 5 mA/cm^2 is applied. (pMOSFET, Tox = 45 Å, W/L = 10/0.25 μm, J_{stress} = 100 mA/cm^2) 90

Fig. 5.6 Associated I-V characteristics of oxide at various stages of electrical stressing as shown inset. Reverse bias anneal applied after onset of QB. It can be seen that gate leakage recovers back to fresh after 500 s of reverse bias anneal. Sample used is the same as Fig. 5.5. 91

Fig. 5.7 DCIV spectra for p-MOSFET oxide at various stages of electrical stressing. The sample used is the same as Fig. 5.8. 92

Fig. 5.8 Associated I-V characteristics at various stages of constant current stressing. After quasi-breakdown is attained, same polarity stressing is continued before application of a reverse bias stress. Gate leakage current after application of reverse bias shows reduction till SILC level. (T_{ox} = 45 Å, W/L = 10/0.5 μm, p-MOSFET). 93

Fig. 5.9 Quantitative DCIV spectra measurement at various stages of current stressing. Peak recombination current I_B is related to interface traps while the lateral shift of peak I_B is related to oxide bulk traps. At reverse bias anneal, oxide bulk traps recover to initial values at fresh state while interface traps remains unchanged. 93

Fig. 5.10 Variation of gate voltage under bipolar constant current density J_{stress} = +/- 10 mA/cm^2 at pre-QB and post-QB with successive alternating stressing and bias annealing. (T_{ox} = 45Å, W/L = 10/0.2 μm, n-MOSFET) 94

Fig. 5.11 DCIV spectra of p-MOSFET stressed to QB. Post-QB thermal annealing performed at various temperatures as shown inset. The thin line linking up the maxima of the I_B spectra reflects the level of oxide bulk traps during SILC while the thick line reflect bulk trap level due to the thermal annealing. It can be observed that thermal anneal results in both a positive shift in the spectra and reduction in the maxima of

I_B showing reduction in interface traps and bulk traps.(T_{ox} = 45 Å, J_{stress} = 50 mA/cm^2, W/L = 10/1.0 μm, p-MOSFET). 95

Fig. 5.12 Gate current leakage current after QB and with post-QB thermal annealing at successively higher temperature for 10mins each. (T_{ox} = 45 Å, J_{stress} = 50 mA/cm^2, W/L = 10/1.0 μm, p-MOSFET). 96

Fig. 5.13 DCIV spectra of p-MOSFET stressed to QB. Post-QB thermal annealing carried out at 200^{o}C for varying period of annealing durations from 5 mins to 200 mins. (T_{ox} = 45 Å, J_{stress} = 20 mA/cm^2, W/L = 10/1.0 μm, p-MOSFET). 97

Fig. 5.14 Gate leakage current at QB and after post-QB thermal annealing treatment at 200^{o}C for varying period of time as shown inset. 98

Fig. 5.15 Carrier separation measurement of source current component (holes current) for p-MOSFET under successive thermal anneal. Sample used is the same as Fig. 5.14 98

Fig. 5.16 Carrier separation measurement of well current component (electron current) for p-MOSFET under inversion conditions after successive thermal anneal. Sample used is the same as Fig. 5.14 99

Fig. 5.17 Gate I-V characteristics for p-MSOFET stressed till unrecoverable QB with post-QB thermal annealing at 200^{o}C for varying period of durations as shown inset. (T_{ox} = 45 Å, J_{stress} = 50 mA/cm^2, W/L = 10/1 μm, p-MOSFET) 100

Fig. 5.18 Carrier separation for source I-V characteristics on the same p-MSOFET used in Fig. 5.17 with post-QB thermal annealing treatment. 100

Fig. 5.19 Carrier separation measurement showing substrate I-V characteristics for p-MOSFET used in Fig. 5.17 with post-QB thermal annealing treatment. 101

Fig. 5.20 Gate leakage current at fresh stage and after electrical stresses. Oxide was stressed till onset of QB as shown by initial QB and continual stressing result in QB(2) before being subjected to reverse bias and thermal anneal. The 'recovered' oxide was then subjected to additional electrical stress till second QB as shown by QB(3). 102

Fig. 5.21 Associated DCIV spectra of oxide stressed till QB and subjected to bias and thermal anneal. It can be observed that the combined effects of bias and thermal anneal results mainly in a positive shift in the DCIV spectra of the post-QB oxide and reduces I_B to a lower level respectively. 103

Fig. 5.22 Evolution of the gate voltage of oxide subjected to CCS till QB as shown in (a) and subsequently subjected to bias and thermal anneal. (b) shows the subsequent evolution of the gate voltage when the 'recovered' oxide is re-subjected to CCS till a second QB. 103

Fig. 5.23 (1) Microscopic model of hole trapping (A)-(B) forming an E' center and detrapping (C) along with charge compensation and bond reformation as proposed

by Lelis *et al.* [5.13] (2) Two spatial equivalent trap levels that electrons can tunnel to, corresponding to the ground and excited state of the E' center. (After [5.13]) 105

Fig. 5.24 Schematic diagram illustrating (A) Reverse bias annealing (B) Thermal annealing without bias. With bias anneal, both levels of trapped holes can be annealed while in thermal annealing, hole trap with energy level above Si conduction band requires electrons with energy above conduction band to be deactivated. 106

Fig. 5.25 (a) Proposed localized trap region (LTR) model: hole trapping predominantly at anode causes an energy band distortion and results in F-N conduction for electrons and direct tunneling for holes at the recoverable QB. (b) Thermal annealing results in electron-hole pair compensation and the reduction of localized trap region (LTR). The shaded regions represent bandgap narrowing due to formation of LTR while the thick lines show the resultant oxide energy band. 107

Fig. 6.1 HRTEM cross section of 13 Å gate oxide (left) and C-V measurements and simulation results (solid lines) fitting to 13 Å oxide thickness by Berkeley QMCV modeling (right). 116

Fig. 6.2 Evolution of gate voltage under constant current stressing (J_{stress} = -50 mA/cm^2) with gate injection for thick (45 Å) and thin (20 Å) gate oxide. (T_{ox} = 45 Å & 20 Å, Channel Area = 10 μm^2, p-MOSFET). 117

Fig. 6.3 (a) Evolution of gate voltage under CCS for thick (45 Å) p-MOSFET oxide after onset of QB. (b) Associated DCIV spectra measured at interval specified in (a). Base recombination current increases continuously under stressing till QB. At QB, DCIV spectra observed to overlap with no further increases. (Channel Area = 10 μm^2, p-MOSFET). 119

Fig. 6.4 Comparison of DCIV spectra for (a) thick (45 Å)and (b) thin oxide (20 Å) p-MOSFET under constant current stressing till QBs. (a) For thick oxide (45 Å), base recombination current increases continuously under stressing till QB. At QB, DCIV spectra observed to overlap with no further increases. (b) For thin oxide (20 Å), DCIV spectra increases even after QB with step-like increases in correlation with gate leakage current. (Locus of $I_{B,max}$ and $V_{g,max}$ is shown by the dotted lines). (Channel Area = 10 μm^2, p-MOSFET). 120

Fig. 6.5 Correlation of gate leakage current I_g and base recombination current I_B, which directly reflect interface trap density N_{IT} under constant voltage stressing for (a) small area samples (W/L = 20/0.5 μm) and (b) large area samples (W/L = 50/50 μm). Inset figure shows the percentage change in I_g and I_B. (Tox = 20 Å, p-channel MOSFET). 122

Fig. 6.6 Correlation of gate leakage current I_g and peak base recombination current $I_{B,max}$ under constant voltage stressing for substrate injection. Similar to negative gate bias (gate injection), gate leakage current observed to bear one-to-one correspondence to

base recombination current which is directly correlated to interface traps. (T_{ox} = 20 Å, W/L = 20/0.5 μm, p-MOSFET). 123

Fig. 6.7 Percentage change of gate leakage current (I_g–$I_{g,initial}/I_{g,initial}$) for different gate voltage under constant voltage stressing. The spike for ΔI_g at gate voltage near to zero volts is due to background noise. Unlike ITT, ITET occurs throughout the entire voltage measurement range. (T_{ox} = 20 Å, Area = 10 μm^2, p-channel MOSFET). 125

Fig. 6.8 Carrier separation showing holes (shown by source current I_s) and electrons current (shown by well current I_w) for small channel area p-MOSFET stressed under negative constant gate voltage. (T_{ox} = 20 Å, Area = 10 μm^2, p-channel MOSFET). 125

Fig. 6.9 (a) Discrete gate current density increase for different channel areas (b) normalized gate current increase under constant voltage stress for different channel area ranging from 10 to 2500 μm^2. (T_{ox} = 20 Å, p-MOSFET). 127

Fig. 6.10 Carrier separation characteristics for 13 Å gate oxide in both depletion and accumulation. Source, n-well and gate current indicated by I_s, I_w and I_g respectively. Drain electrode is not connected. (Channel Area = 100 μm^2, p-MOSFET). 128

Fig. 6.11 (a) Evolution of current-voltage characteristics at fresh and after constant voltage stressing (V_{stress} = -3.0 V). (b) Associated I-V characteristics at onset of progressive BD and post PBDs. (T_{ox} = 13.4 Å, Channel Area = 10 μm^2, p-MOSFET). 130

Fig. 6.12 Evolution of current-voltage characteristics at fresh and after constant voltage stressing (V_{stress} = + 3.0 V). (T_{ox} = 13.4 Å, Channel Area = 10 μm^2, p-MOSFET). 131

Fig. 6.13 (a) Percentage change in gate leakage current under different stressing gate voltage in the initial stage of PBDs. (b) In the subsequent PBDs stages, leakage current proportional to logarithmic of stressing time. It can be observed that gate leakage current follow a power relation with stressing time in the initial stage (a) and a linear logarithmic time dependence as shown inset in the subsequent stage (b). 132

Fig. 6.14 Graph showing gate leakage current time dependence parameters A and voltage dependence parameter, B as defined in (1) for 2nd stage of PBDs versus stressing gate voltage. (Channel Area = 100 μm^2, p-MOSFET). 134

Fig. 6.15 (c) Percentage increase in gate leakage current after onset of 2nd stage PBDs for different sample areas. 136

Fig. 6.16 Degradation rate or defect generation rate defined by Jg = Pg * Qinj where Pg is the defect generation / gate degradation rate as shown for second stage of PBDs. It can be observed that gate degradation increase as sample area decreases with eventual saturation at very small area. 137

Fig. 6.17 Lifetime projection versus stressing gate voltage using 100% increase in gate leakage current as failure criterion. Lifetime for 100% increase in gate leakage $t_{100\%}$

$_{Ig}$ is much shorter than the conventional time-to-complete breakdown t_{BD}. (T_{ox} = 13.4 Å, p-MOSFET). 138

Fig. 6.18 Evolution of normalized gate leakage current $(I_g - I_{g,0})/I_{g,0}$ on a 13.4 Å gate oxide when stressed under constant voltage stress (as shown by the solid symbol). Using (5.6), the cumulative frequency of localized spots occurrence F(t) can be determined and it's associated Weibull plot (with various values of Rj) versus logarithmic of stressing time is as shown (shown by the various open symbols). It can be observed that the derived Weibull plots can be separated into 2 regions, A & B. The derived Weibull shows a good linearity when Rj = 3 for both region A and B, deviating only at the extreme short and long stressing time due to censoring effects. 142

Fig. 7.1 Energy band diagram and tunneling current components for p-MOSFET with metal gate under (a) inversion (negative gate bias) and (b) accumulation (positive gate bias) conditions. The dominant components of gate currents under both polarities are the ones which tunnel through the IL: Js under $-V_g$ and Jw under $+V_g$. 151

Fig. 7.2 J_g-V_g characteristics of p-MOSFETs with HfAlO dielectrics. Source and n-well are grounded. The dominant components of gate currents under both polarities are the ones which tunnel through the IL: J_s under $-V_g$ and J_w under $+V_g$. 152

Fig. 7.3 Charge trapping characteristics (a) under positive V_g, and (b) negative V_g. Positive bias stress causes electron trapping, while negative bias stress results in hole trapping for $|V_g| > -3$ V. Sample area is 100 x 100 μm^2. 155

Fig. 7.4 Inversion capacitance-voltage curves of p-MOSFETs before and after stress under (a) negative gate bias and (b) positive gate bias for a period of 1000 s and 2000 s. A flatband voltage shift to the left after negative bias stress shows positive charge trapping within the dielectric and IL. 156

Fig. 7.5 (a) Evolution of gate leakage current in p-MOSC and (b) evolution of current-voltage characteristics under negative constant voltage stress. It can be observed that there are at least 2 distinct stages of breakdown as shown by pBD1, pBD2 and pBD3. Area of sample used is 100 x 100 μm^2. 159

Fig. 7.6 (a) Evolution of gate leakage current in p-MOSC and (b) evolution of current-voltage characteristics under positive constant voltage stress. It can be observed that there are only 1 distinct stages of breakdown. Area of sample used is 100 x 100 μm^2. 160

Fig. 7.7 (a) Relative changes of Js and Jw currents during negative CVS (V_g = -3.0 V) on p-MOSFET. (b) Identical data with (a) but plotted in wider scale. The high-K bulk breakdown happens first at the initial stage of breakdown. 162

Fig. 7.8 Associated (J-V characteristics) carrier separation measurement of Fig. 6.8, showing both source and well current component before and after bulk BD and IL BD. It can be observed that at the first onset of Bulk BD (shown by shaded symbols), J_w

increases significantly throughout the entire range of gate voltage while J_s, which reflect the IL condition, increases only marginally. Subsequent occurrence of IL BD (shown by solid symbols), result in significant increase in both J_s and J_w.....163

Fig. 7.9 (a) Relative changes of J_s and J_w currents during positive CVS (V_g = +3.2 V) on p-MOSFET. Only interfacial layer breakdown observed after about 470s of stressing, shown by the larger increase in J_w. (b) Associated I-V characteristics after positive CVS but measured under both negative and positive gate voltage regime. The IL leakage currents have increased by orders of magnitude, while the bulk leakage current did not increase significantly due to the intact high-*K* bulk layer. 165

Fig. 7.10 (a) Relative changes of Js and Jw currents during positive CVS (Vg = 3.2 V) on n-channel MOSFET and (b) its associated J-V characteristics. The dominant electron current, J_s, increase faster than that of the subservient well current, J_w, which reflect the valence electron current. At onset of breakdown after stressing for 15 s, interfacial layer breaks down, which translate to much higher leakage current for Js while the well leakage current J_w which tunnel through a thicker portion of the high-*K* layer is less affected. 167

Fig. 7.11 Weibull distributions for V_{BD} under ramped gate voltage sweep for p-channel MOS capacitors (shown by shaded symbols) and n-channel MOS capacitors (shown by open symbols) with source/drain implant. The Weibull slope for breakdown voltage V_{BD} under $-V_g$ sweep is much steeper than that for $+V_g$ sweep for both n-channel and p-channel MOS. 169

Fig. 7.12 Weibull distributions for charge-to-breakdown (QBD) (a) under negative constant voltage stresses and (b) under positive constant voltage stresses. It is observed that Weibull distribution for $+V_g$ CVS deviates from the linear line at low QBD for higher gate bias stressing due to temporal resolution of measurement setup, especially for early failure devices with low time-to-breakdown. High β (Weibull slope) value under $-V_g$ and low β value under $+V_g$ are observed. Sample area is 100 x 100 μm^2. 172

Fig. 7.13 Weibull distributions for charge-to-breakdown (Q_{BD}) on p-channel MOS capacitors under (a) negative constant voltage stresses and (b) under positive constant voltage stresses. Reasonably high β (Weibull slope) value under $-V_g$ CVS and low β value under $+V_g$ CVS are observed. Sample area is 10^{-4} cm^2. 173

Fig. 7.14 A schematic drawing for a breakdown model using charge trappings at different spatial locations in high-*K*/IL stack dielectric with a metal gate structure. (a) For $-V_g$, electron trapping occurs mainly in the bulk while hole trapping occurs near to the IL. The columbic force of the trapped charges distorts the energy band diagram, leading to a preferential breakdown in the bulk. (b) For $+V_g$, only electron trapping occurs. The band distortion in the high-K bulk would be smaller, leading to higher possibilities of interfacial layer breakdown. 174

List of Tables

Table 1.1 Selected data from latest ITRS 2003 update (After [1.2]) ……. 3

Table 2.2 Summary of quasi-breakdown conduction mechanism …….……. 19

Table 6.1 A summary of criteria for determining the dominant breakdown mechanism for all the possible combinations of gate bias in both n- and p- MOSFETs …….. 166

Chapter 1

Introduction

1.1 Dimension Scaling and Future Trends of Microelectronics

Microelectronics is becoming an important and integral part of modern living. It is interesting to note that in almost every part of our lives, including medical, transport, entertainment, communication and military defense, electronics is invariably present. The increasing miniaturization of electronics to even smaller sizes, through device scaling, novel process fabrication and device structures to its ultimate limit - nano dimensions, introduces changes, which were previously unimaginable. Nanoscience engineering provides new knowledge and capability to design and build materials at atomic scale. Yet these changes require tremendous engineering ingenuity, and researches into such new materials and its underlying science, are imperative.

Integral to the entire electronics chip is the transistor. The first oxidized silicon metal-oxide-semiconductor field effect transistor (MOSFET) was first proposed and fabricated by Kahng and Attalla in 1960. Since then, the inherent structure of the transistor has remained almost unchanged till today. In addition, due to the various benefits of silicon dioxide as the gate insulator, there has been little or virtually no change in the gate dielectric material. Ever since then, the technological advancement in electronic circuitry is achieved merely by reducing device dimensions to achieve higher speed and higher packing density. Decades of continuous technological improvements in CMOS technology have made it the present dominant Very Large Scale Integration (VLSI) technology. Beneficial results from such intense scaling can be observed in bit-density increase, speed/performance as well as reliability improvement and defect reduction resulting in significant yield improvement. While

Moore's law[1] has been able to predict the dimension scaling in the microchip very well over the past decades, there are increasing evidences that certain fundamental barriers will be approached, which may limit the continuous phenomenal growth in transistors' density [1.1].

Figure 1.1 shows the device scaling for the last thirty years. It can be seen that as device channel length is aggressively scaled downward, gate oxide thickness is also scaled to avoid short channel effect and to maintain drive current capability. There are, however, increasing evidences that oxide scaling may be reaching a limit due to the tradeoff in gate leakage and oxide reliability for ultra-thin oxides.

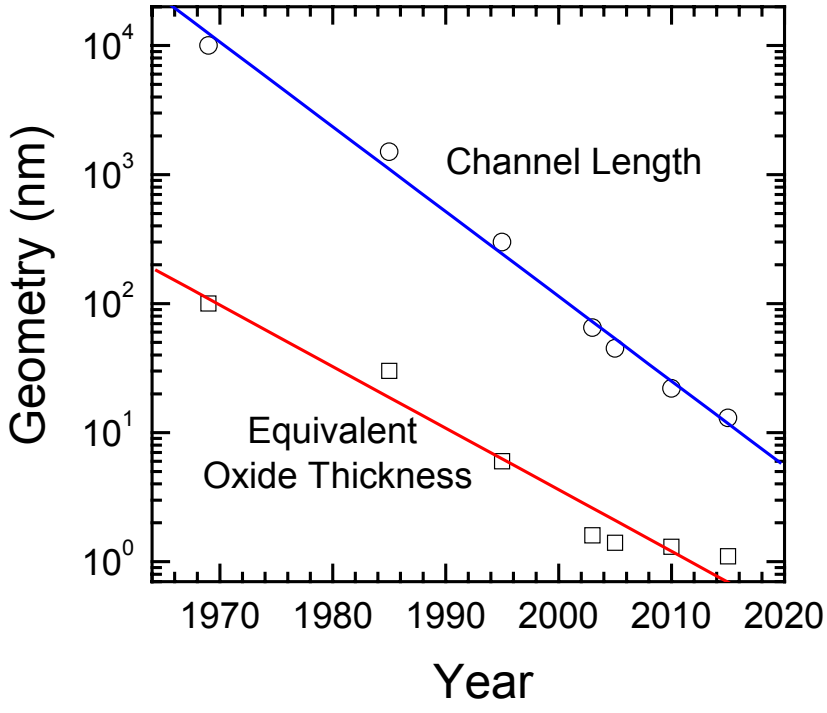


Fig. 1.1 Geometry scaling for MOSFET device channel length and equivalent oxide thickness for low power application MOSFETs. Equivalent oxide thickness (EOT) is used instead of physical oxide thickness due to the potential change in dielectrics to high-*K* material for 65 nm technology node.

Table 1.1 in the next page, shows the technological roadmap for the semiconductor industry in the coming 10 years [1.2]. It can be observed that there are a few significant issues that do not have any solutions presently. For accelerated

[1] The observation was made in 1965 by Gordon Moore, who found that the number of transistors in integrated circuits expressed in per square inch, has doubled every year since the integrated circuit was invented. Recently, the pace has slowed down a bit, with data density doubling approximately every 18 months instead of 12 months.

MOSFET gate length scaling to continue, the following key issues shown below have to be addressed.

- Accelerated need for high-*K* gate dielectric solution for dealing with increased MOSFET gate leakage
- Accelerated need for dual metal gate electrodes and next generation contact solutions due to incompatibility of polysilicon with high-*K* dielectrics and poly depletion effect
- Accelerated need for ultra-shallow highly activated extensions

In particular, gate dielectric using gate oxides will face significant challenges as gate oxide thickness approaches the direct tunneling regime of below 30 Å, requiring rapid supply voltage derating as shown in Table 1.1. Enhanced direct

Table 1.1 Selected data from latest ITRS 2003 update (After [1.2])

Year of Introduction " Technology Node"	2003	2004 90nm	2005	2006	2007 65nm	2009	2010 45nm	2013 32nm
Physical gate Length (nm) for low operating power	65	53	45	37	32	25	22	16
Minimum Supply Voltage (volts)	1	0.9	0.9	9	0.8	0.8	0.7	0.5 - 0.6
Equivalent Physical Oxide Thickness (nm)	1.6	1.5	1.4	1.3	1.2	1	0.9	0.8
Gate Dielectric Leakage @ 25 deg C (A/cm^2)	0.51	1.89	2.22	2.7	5.21	6.67	11	21
Contact maximium resistivity ($x\ 10^{-7}$ ohms-cm^2)	1.93	1.62	1.44	1.2	1.05	0.72	0.61	0.171
Drain Extension Xj (nm)	24.8	20.4	17.6		13.8	8	7.2	10.4

Solution Exist Solution being pursued No known solution

tunneling leakage current due to quantum-mechanical (QM) tunneling probabilities of electrons results in higher standby leakage current I_{off} and anomalous capacitance-voltage behavior that progressively destroys transistor operation characteristics. This places a theoretical limit on the usage of SiO_2, which has leakage current in the excess of 5 A/cm^2 for the 13Å thickness regime [1.2]. Using electron energy loss spectroscopy (EELS), it is observed that the two interfacial layers overlap when SiO_2 layer thinner than 13 Å is used [1.3]. At this thickness regime, gate leakage current

becomes very large ~ 10^2 A/cm^2 and the insulating nature of SiO_2 is almost completely lost. As a result, it is obvious that for gate dielectrics with equivalent oxide thickness of 13Å and below, other materials such as high-*K* gate dielectrics will be required [1.3],[1.4].

Besides the excessive gate leakage observed in ultra-thin gate dioxides, other hosts of problems also arise from this frantic device scaling. In particular, reliability has currently become an important issue due to several factors. Firstly, although device dimensions are scaled downward, the applied voltage cannot be scaled proportionately, due to the presence of a large mixture of logic/digital and input/output devices in a single chip which have different power requirement. This non-proportional voltage scaling has resulted in increasing electrical field which is very detrimental to device lifetime. Higher field has led to increased leakage current, power dissipation and enhanced device temperature, both of which have very adverse effects on device operation. Fig. 1.2 shows the long term reliability requirement for the MOSFET device. By the year 2010, long term reliability requirement of a transistor may need to be lower than 1 Failure-in-Time (FIT). This requirement is

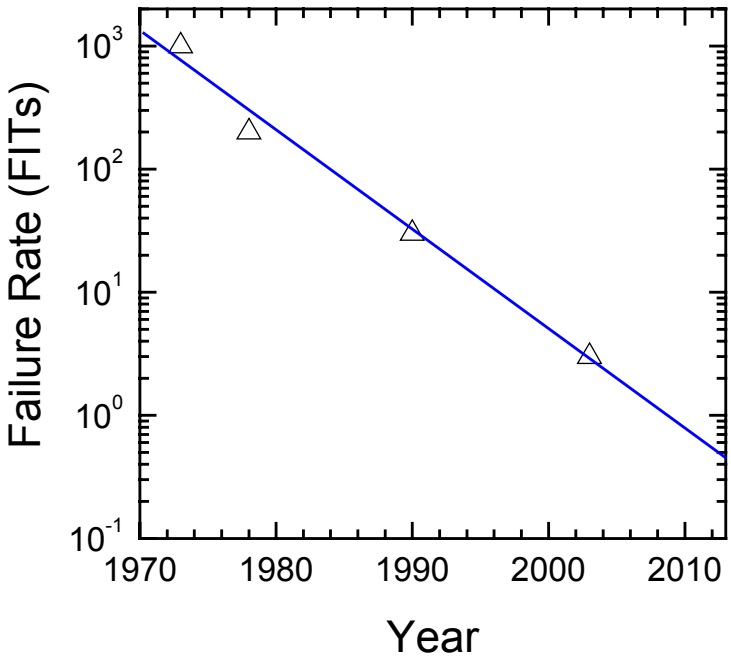


Fig. 1.2 Trend in long term reliability requirement for MOSFETs. 1 FITs = 1 failure per 10^9 device hours.

even more difficult to achieve, considering that new high-*K* dielectrics will be needed by this time. While many materials including metal oxides such as Ta_2O_5 and TiO_2, Y_2O_3, HfO_2 and silicates such as Zr silicate have been proposed, they are generally not thermally stable on silicon and the formation of SiO_2 and metal silicides often occurs at the interface [1.5] This decreases the effective dielectric constant and hence its capacitive effect. Other issues concerning high-*K* dielectrics include mobility degradation, boron penetration, thermal stability, high fixed charge density Q_f and high gate leakage current [1.5].

The introduction of new materials such as high-*K* gate dielectrics and metal gate electrodes also introduces other problems in terms of device reliability, process integration and new types of defect generation and detection. The degradation and breakdown mechanism for future high-*K* dielectrics is presently unclear and may become a potential barrier to successful implementation of high-*K* dielectrics. These issues will become even more critical considering the rapid changes in materials needed to keep pace with dimension scaling requirements. In view of the numerous challenges facing high-*K* dielectrics, researchers have tried to prolong the usability of SiO_2 by incorporating nitrogen to enhance its *K* value. In this respect, ultra-thin oxynitride dielectrics (T_{ox} ~ 14 Å), has been fabricated with very good device characteristics [1.6]. Beyond this thickness, however, the fundamental limit of SiO_2 will still be reached and implementation of high-*K* dielectrics becomes unavoidable.

1.2 Summary and Motivation of Thesis

As gate dielectrics scales downwards, various reliability issues have surfaced. For gate oxides thinner than 50Å, a phenomenon known as quasi-breakdown (QB) [1.7], was observed prior to complete breakdown. While extensive studies have been conducted, the conduction mechanism of quasi-breakdown remains controversial. Two main models that are widely cited are the direct tunneling [1.7] and percolation path [1.8] models. The full understanding of quasi-breakdown mechanism and its conduction kinetics will be one of the main focuses and objectives in this thesis.

As ultra-thin oxides below 20 Å are used for the 130 nm technology node and beyond, various other degradation mechanisms are also observed in such ultra-thin

oxides. Progressive breakdown leading to a progressive increase in gate leakage current is observed in these ultra-thin oxides [1.9]. The degradation mechanism, modeling and reliability extrapolation in such thickness regime are necessary for commercial implementation of such ultra-thin oxides and will be studied in chapter five of this thesis.

Beyond the 65 nm technology node, the International Technology Roadmap for Semiconductors (ITRS) 2003 shows that high-*K* dielectrics will be required if the current planar single-gated transistor structure is to be continued. Double or even triple-gated structure including FINFETs [1.10] and vertical transistor may mitigate this requirement but the process complexity of such structures may be too intimidating for the near-term implementation. As such, the intrinsic reliability of high-*K* dielectrics is important for successful implementation and integration with future CMOS processes. In addition, while there have been a substantial study in the reliability of high-*K* dielectrics stack, the current methodologies available are tedious and involves large amounts of devices testing and sampling. This is due to the inherently low Weibull slope [1.11] observed in high-*K* dielectrics, which results in significant scattering in the breakdown distribution statistics. Hence one of the objectives in this thesis is to provide a simple and direct methodology of studying and characterizing the reliability of high-*K* dielectrics stacks. A novel carrier separation technique is proposed which can effectively distinguish the bulk layer or interfacial layer initiated breakdown.

1.3 Thesis Outline and Original Contributions

This thesis consists of eight chapters and is arranged as follows. Chapter one describes the CMOS scaling and the accompanying issues. Chapter two describes the various oxide degradation mechanisms observed for various thickness regimes. Various reliability issues regarding oxide degradations [1.12] are raised and these will be addressed in the following chapters. Chapter two describes the two experimental measurement techniques: direct current current-voltage (DCIV) method and carrier separation, which will be used throughout the thesis. The underlying principles and a comparative study with other commonly used techniques is presented. Chapter four

analyzes the effect of bipolar current stressing on 45 Å thick SiO_2. It is observed that charge-to-quasi-breakdown Q_{QB} is not the same for bipolar and unipolar current stressing. In particular, it is observed that on small sample areas (< 3 μm^2), bipolar current stressing results in a lower Q_{QB} than both positive and negative unipolar current stressing. In larger sample areas (≥ 100 μm^2), however, Q_{QB} for bipolar current stressing is similar to negative unipolar current stressing (gate electron injection). The result suggests that trap generation is not uniformly distributed and bipolar Q_{QB} is strongly dependent on sample channel area. Using bipolar current stressing, it is also observed for the first time, that QB can be separated into two stages – recoverable and unrecoverable QB, which are characterized by its electrical recoverability. By the applying carrier separation technique to the two stages within QB, it is observed that within recoverable QB, there exist the possibilities of either hole dominance or electron dominance. Using the Fowler-Nordheim (F-N) equation, it is further observed that the electron leakage current at QB can be adequately described by the Fowler-Nordheim tunneling equation. The hole leakage current, on the other hand, follows a direct tunneling equation. By using a simple model of holes trapping at the anode, the various QB phenomena can be explained and the initial locally physical damage region (LPDR) model [1.7] which was earlier proposed by Lee and Cho *et al.*, is further ascertained.

In chapter five, the annealing behavior of post-QB oxide under thermal and electrical bias anneal is described. It is observed that reverse bias anneal is able to detrap the positive charges within the oxides, thus lowering the QB leakage current back to the stress-induced leakage current (SILC) level. Under thermal annealing, it is observed that substrate current (holes) can be reduced to pre-stress levels while well leakage current (electron) saturates above a certain level. The result suggests distinctive trap levels for electron and hole conduction in post-QB oxides and supports the earlier trap-induced QB breakdown model in chapter three.

In chapter six, oxide degradation for ultra-thin oxide (< 20 Å) is described. It is observed that unlike thicker oxide, QB in ultra-thin oxide (20 Å) can be directly correlated to discrete increases in interfacial traps. Moreover as thickness of silicon dioxide reduces even further till 14 Å, QB is not distinctly observed. Instead, gate

leakage current increases progressively after certain period of stressing. Using a localized multiple breakdown spots model, it is shown that the gate leakage can be attributed to a multiple occurrence of breakdown [1.13]. A new failure criterion based on gate leakage current density is proposed and this is shown to be far more practical for ultra-thin oxide than the excessively optimistic conventional time-to-complete breakdown.

As gate silicon dioxides outlive its usefulness at around 14 Å, beyond which direct tunneling leakage current will be too high for general device applications, high-*K* dielectrics is needed for the 65 nm technology node. Chapter seven studies the reliability of high-*K* stacks using a novel carrier separation method. A time-to-breakdown with polarity dependence, is observed under constant voltage stress and this is attributed to breakdown at different layers within the high-*K* stacks [1.14]. Finally, chapter eight concludes the thesis with some suggestions for future research based on the findings and conclusions arrived in this thesis.

Reference

[1.1] Y. Nishi and J. W. McPherson, "Impact of New Materials, Changes in Physics and Continued ULSI Scaling on Failure Mechanism and Analysis," *7th Proc. IPFA* (Keynote Address), pp. 1-8, 1999.

[1.2] International Technology Roadmap for Semiconductors (ITRS) 2003, Semiconductor Industry Assoc., San Jose, CA., Available: http://public.itrs.net/Files/2003ITRS/Home2003.htm.

[1.3] M. L. Green, E. P. Gusev, R. Degraeve, and E. L. Garfunkel, "Ultrathin (< 4nm) SiO_2 and Si-O-N gate dielectric layers for silicon microelectronics: understanding the processing, structure, and physical and electrical limits," *J. Appl. Phys.,* vol. 90, pp. 2057-2121, 2001.

[1.4] D. A. Buchanan, "Scaling the gate dielectric: materials, integration, and reliability," *IBM J. Res. Develop.,* vol. 43, pp. 245-264, 1999.

[1.5] G. D. Wilk, R. M. Wallace, and J. M. Anthony, "High-*K* gate dielectrics : current status and material properities considerations," *J. Appl. Phys*., vol. 89, pp. 5243-5275, 2001.

[1.6] S. Song, J. H. Yi, W. S. Kim, J. S. Lee, K. Fujihara, H. K. Kang, J. T. Moon, and M. Y. Lee, "CMOS Device Scaling Beyond 100nm," *IEDM Tech. Dig.,* pp.235-238, 2000.

[1.7] S. H. Lee, B. J. Cho, J. C. Kim, and S. H. Choi, "Quasi-breakdown of ultrathin gate oxide under high field stress," *IEDM Tech. Dig*., pp. 605-608, 1994.

[1.8] M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, "Soft breakdown in ultrathin gate oxides: correlation with percolation theory of non-linear conductors," *Appl. Phys. Lett.,* vol. 73, pp. 514-516, 1998.

[1.9] Y. Wu, D. Bang, G. Lucovsky, and M. R. Lin, "Time Dependent Dielectric Wearout (TDDW) Technique for Relaibility of Ultrathin Gate Oxides," *IEEE Electron Dev. Lett.,* vol. 20, no. 6, pp. 262-264, 1999.

[1.10] X. Huang, W. C. Lee, C. Kuo, D. Hisamoto, L. Chang, J. Kedzierski, E. Anderson, H. Takeuchi, Y. K. Choi, K. Asano, V. Subramanian, T. J. King,

J. Bokor, and C. Hu, "Sub-50-nm FINFET : PMOS," *IEDM Tech. Dig.,* pp. 67-70, 1999.

[1.11] T. Kauerauf, R. Degraeve, E. Cartier, C. Soens, and G. Groeseneken, "Low Weibull Slope of Breakdown Distributions in High-*K* layers," *IEEE Electron Dev. Lett.*, vol. 23, pp. 215-217, 2002.

[1.12] D. J. Maria, E. Cartier and D. Arnold, "Impact Ionization, trap creation and breakdown in silicon dioxide films on silicon," *J. Appl. Phys*., vol. 73, p. 3367, 1993.

[1.13] W. Y. Loh, B. J. Cho, M. F. Li, Daniel S. H. Chan, C. H. Ang, J. Z. Zhen, and D. L. Kwong, "Localized Oxide Degradatioin in Ultra-Thin Gate Dielectric and its Statistical Analysis," *IEEE Tran. Electron Dev.,* vol. 50, pp. 967 – 972, 2003.

[1.14] W. Y. Loh, B. J. Cho, M. S. Joo, M. F. Li, S. Matthew, Daniel S. H. Chan, and D. L. Kwong, "Analysis of Charge Trapping and Breakdown Mechanism in High-*K* Dielectrics with Metal Gate Electrode using Carrier Separation," *IEDM Tech. Dig.,* pp. 927-930, 2003.

Chapter 2

Literature Review: Gate Dielectric Degradation

2.1 Impact of device scaling on gate dielectric degradation

Aggressive continual device scaling has resulted in reduction of device dimensions without proportional reduction in supply voltage. This has led to high field effect phenomena such as hot carrier degradation. At channel electric field above 4 x 10^4 V/cm (corresponds to V_g = 2V, V_d = 5 V, T_{OX} = 200 Å), a significant hot-carrier effect occurs which greatly degrade both p- and n-MOSFET [2.1]. Fig. 2.1 shows a hot carrier effect in an n-MOSFET with hot electron injection into the gate oxide near the drain region due to high electric field at that region. Impact ionization and avalanche multiplication occurring near the depletion layer edge, also result in the generation of hot holes and electrons which were then injected into the gate [2.1].

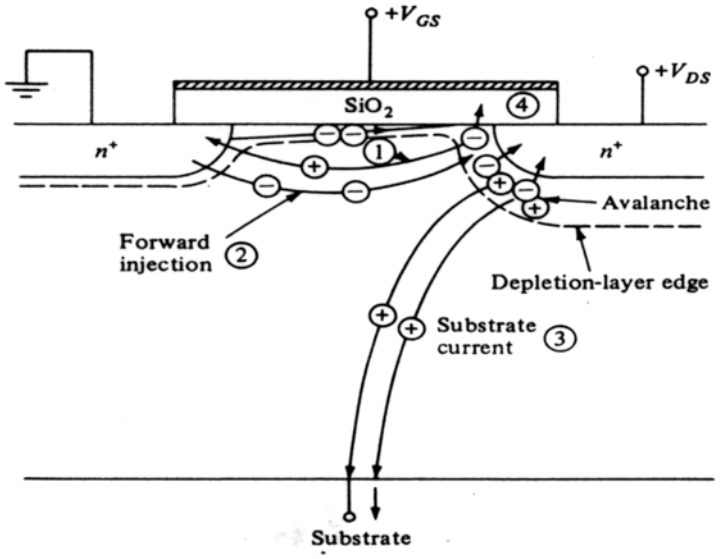


Fig. 2.1 Hot carrier generation and various current components in n-MOSFET. (After [2.1])

As gate oxide is scaled below 100 Å, another phenomenon, known as stress-induced leakage current (SILC) becomes more pronounced. First discovered by

Maserjian *et al.* [2.2], SILC has become a major reliability concern for thin gate oxide due to its increase in gate leakage current. This increased current consumes additional power, which becomes undesirable especially in low-power and portable applications. It was observed that SILC strongly increases as oxide thickness decreases from 100 Å, down to 50 Å. As gate oxide reduces below 50 Å, it was however observed that the relative magnitude of SILC defined by $(J_{g,stress} - J_{g,initial})/J_{g,initial}$ (where $J_{g,initial}$ is the current density), prior to electrical stresssing decreases instead. This is often attributed to reduction in stress induced oxide bulk traps [2.3],[2.4]. Recently, Wu *et al*. have shown that this 'reduction' in SILC is due primarily to SILC effects being overshadowed by the higher direct tunneling current and not the reduction in trap density for the different oxide thickness [2.5]. The conduction mechanism of SILC is generally believed to be due to trap-assisted tunneling through the degraded oxide [2.6]-[2.8]. Using quantum yield studies and carrier separation technique, Takagi *et al.* have shown that inelastic trap- assisted tunneling (TAT) occurs under SILC, with an energy loss of 1.5eV due to energy relaxation of injected electrons at the SiO_2 traps [2.6]. On the other hand, Ielmini *et al.* [2.8] have proposed hole and electron recombination and trap-assisted tunneling (RTAT) as the main conduction mechanism for SILC (Refer to Fig. 2.2c).

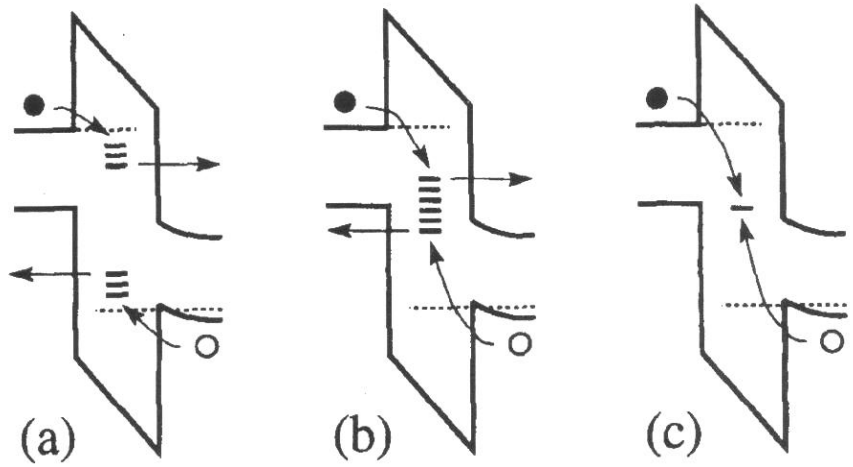
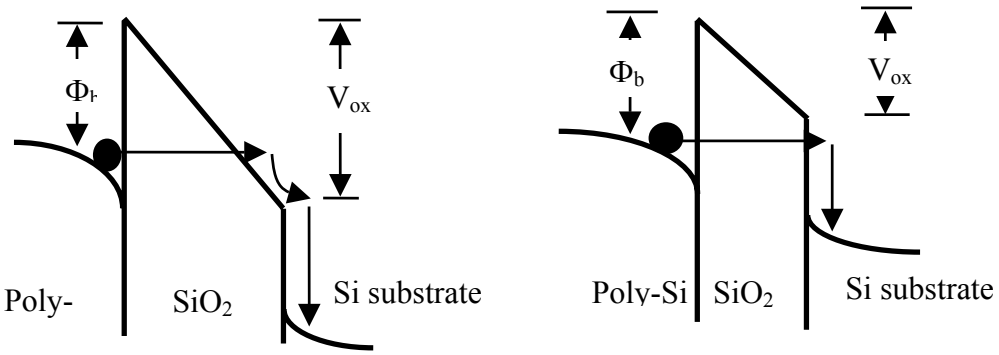


Fig. 2.2 Schematics of three possible conduction mechanisms leading to SILC leakage current. (a) Trap-assisted tunneling (TAT) at distinct defect locations with energy relaxation (b) Trap-assisted tunneling at same defect distribution (c) Tunneling and recombination at oxide defect sites (RTAT). (After [2.8])

Moreover, the charge state of the SILC-related centers is also not clear: Neutral traps have been generally considered, although recent evidences show a clear correlation between SILC and trapped holes [2.9].

For gate oxide thinner than 50 Å, a new phenomenon known as quasi-breakdown (QB) is observed. Using photon emission studies, it was observed that QB is a localized phenomenon with multiple events occurring before complete breakdown [2.10]. The QB breakdown sites have also been found to be uncorrelated to the final complete breakdown spot, suggesting that both phenomena may be independent of each other. The mechanism of QB is the focus of the study here and will be described in greater detail in the next section 2.3.

As gate oxide is aggressively scaled below 35 Å, the high gate leakage current occurs even at low field due to quantum-mechanical tunneling of electrons. At such thickness regime, the conduction mechanism changes from Fowler-Nordheim (F-N) to direct tunneling, resulting in gate leakage which is significantly higher at low field



(a) Fowler-Nordheim Tunneling (b) Direct Tunneling

Fig. 2.3 Illustration of differences between (a) Fowler-Nordheim (FN) and (b) direct tunneling (DT). FN tunneling occurs when $V_{ox} > \Phi_b$ while DT occurs when $V_{ox} < \Phi_b$.

and relatively insensitive to field effect (Refer to Fig. 2.3). Figure 2.4 in the next page, shows the change of conduction mechanism from F-N tunneling to direct tunneling when gate oxide is reduced below 35 Å. In the direct tunneling thickness regime, a slight decrease in gate oxide thickness results in an increase in order of magnitude in the leakage current. This trend will continue as scaling proceeds below

35 Å, due to its exponential dependence of leakage current on oxide thickness [2.11]. Eventually, high-*K* dielectrics will be required although this may be mediated by nitrided silicon oxide film in the short run [2.12],[2.13].

In retrospect, it can be observed that gate oxide thickness scaling and un-proportionate voltage derating have been accompanied by a host of different reliability issues at each technology node. With future incorporation of high-*K* gate dielectrics, it is expected that a different host of problems will be encountered.

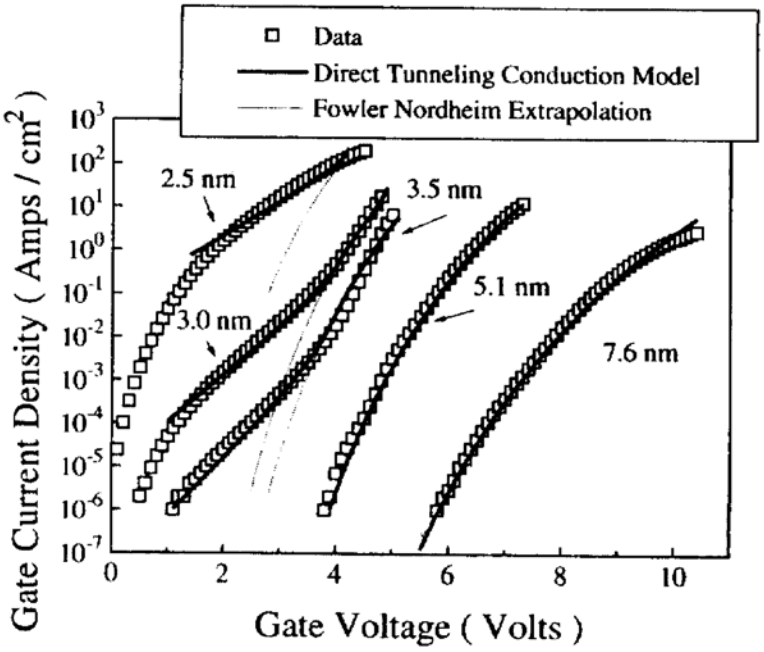


Fig. 2.4 Low-voltage conduction mechanism for thin oxide of various oxide thicknesses. (After [2.11])

2.2 Electrical Stress-Induced Degradation and Breakdown

The degradation of gate oxide under Fowler-Nordheim (F-N) stress can be characterized into 2 main stages – wear out and runaway stage. Since the wear out stage is significantly longer than the runaway stage, which is very rapid, oxide lifetime is dependent on the time for wear out stage to complete. Due to its intrinsic importance, the nature and origin of oxide degradation under F-N stress in the wear out stage have been extensively studied. DiMaria *et al.* have proposed two possible mechanisms for oxide degradation under F-N stress [2.1]. In the first mechanism, electrons with energy greater than 2 eV can release hydrogen from defect sites near the anode interface. The released hydrogen can then diffuse to the cathode-oxide

interface where it creates interface states and oxide electron traps [2.1]. The hydrogen release model (HR) is supported by experimental evidences that oxide containing excess hydrogen has lower charge-to-breakdown Q_{BD}. In the second mechanism, electron with energy greater than 9 eV causes impact ionization at the anode. The holes produced are then re-injected into the oxide due to the electric field and are trapped at deep level traps sites. Recombination of the trapped holes with the injected electrons, results in the formation of interface states and traps near the cathode.

Another widely cited degradation mechanism is the Anode Hole Injection model (AHI), first proposed by Chen *et al.* [2.14]. According to the AHI model, some electrons are trapped in the oxide near the anode. The rest gains enough energy to reach the anode whereby some have sufficient energy to cause impact ionization. The generated holes are then injected back into the oxide, with some trapping in certain localized regions near the cathode. The field enhancement leads to an increase in F-N current through the weak localized spots which degrade further due to positive feedback cycles. In thin oxide, impact ionization cannot occur and instead, the holes are produced when the injected electrons transfer their energy to a valence electron at the anode as shown in Fig. 2.5. Oxide breakdown occurs when a certain level of

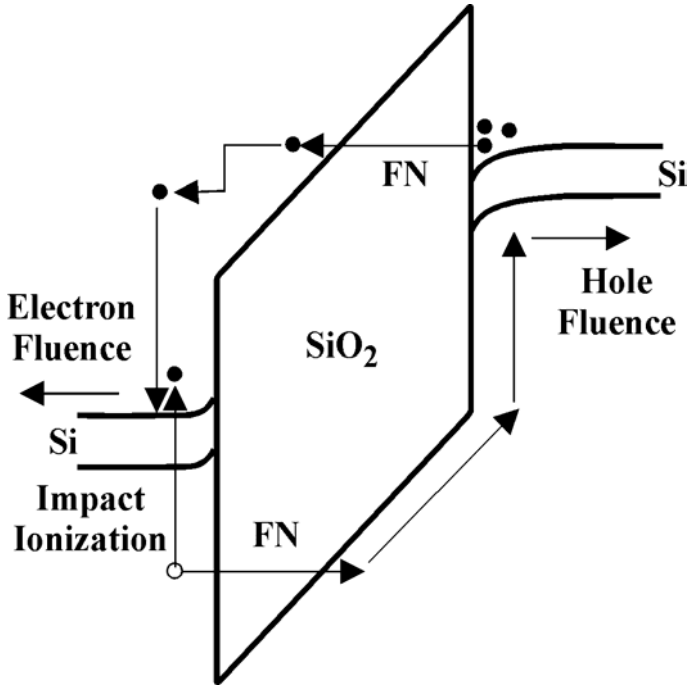


Fig. 2.5 Anode Hole Injection model with an incident electrons arriving at the anode and transferring its energy to a deep level valence electron and in the process, creating a hole which is then injected back into the oxide. (After [2.15])

wearout is reached [2.16],[2.17], and very high leakage current passes through the gate dielectric. In thick oxides, complete breakdown is usually accompanied by high Joule heating resulting in a catastrophic damage to the oxide [2.18].

Based on the electrical degradation mechanism described, early oxide breakdown models for thick oxides can be grouped into thermochemical [2.19] model and 1/E model [2.14] (which is based on the AHI model). In the thermochemical model, electric field interacts with the oxygen-deficient strained Si-Si bonds resulting in dissociation and trap formations [2.20]. Tunneling electrons are not necessary to create the damage which is due primarily to field interactions.

On the other hand, for the AHI model, electron injection creates impact ionization and hole generation which leads to positive enhancement of the internal field and subsequently, to breakdown. Both models are able to explain oxide lifetime at high gate voltage, although the lifetime extrapolation differs widely at low field with the 1/E model being much more optimistic [2.18]. Recently, it is observed that polarity dependence of breakdown exists for ultra-thin oxide in the thickness regime of 40 Å. DiMaria *et al.* proposed that defect generation in the oxide depends on the Fermi level at the anode and is gate voltage driven [2.21]. Based on substrate hot electron injection experiment [2.22], thickness, polarity difference in the Q_{BD} [2.23], and Weibull slope modeling [2.24], it was found that the original (E-model) electric field driven model may not be able to explain the breakdown in ultra-thin oxides. Instead, degradation and breakdown are well described by the release of energy of tunneling electrons at the anode, which is proportional to the applied voltage [2.25].

2.3 Quasi-breakdown Mechanism

While SILC and hot carrier effects are important in oxides with thickness above 50 Å, both effects become mitigated as oxides and voltage reduction results, due to increased device scaling. At the same time, as oxide scales downwards in thickness, conventional complete breakdown becomes less prevalent. Instead, oxide breakdown characterized by quasi-breakdown (QB) [2.10],[2.19]-[2.30], becomes more important and prevalent for oxides less than 50 Å.

Also known as soft breakdown [2.32] or B-mode SILC [2.37], QB is observed as gate oxide thickness goes below 50 Å, approaching the direct tunneling regime. Unlike complete or dielectric breakdown (CB), QB is characterized by gate leakage

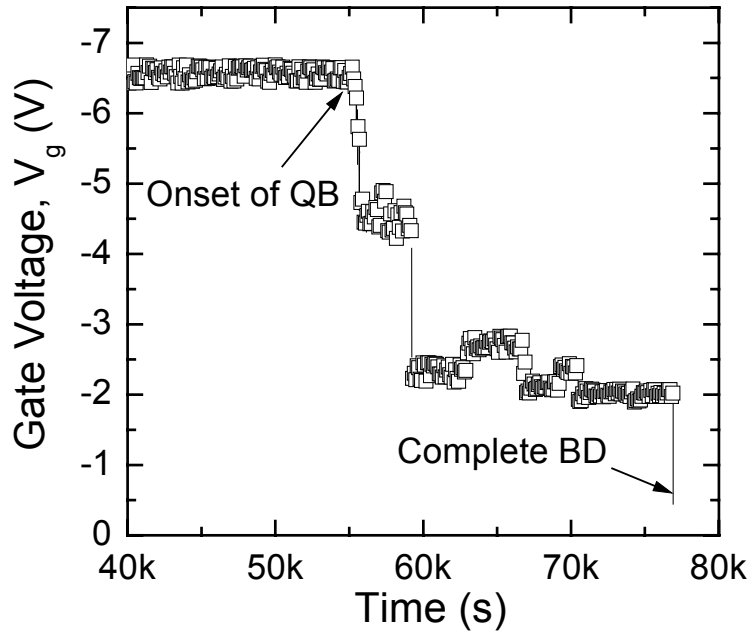


Fig. 2.6 Evolution of gate voltage under constant current stress till complete breakdown. At quasi-breakdown (QB), gate voltage drops due to enhanced leakage path with gate voltage magnitude still significantly higher than at complete breakdown.

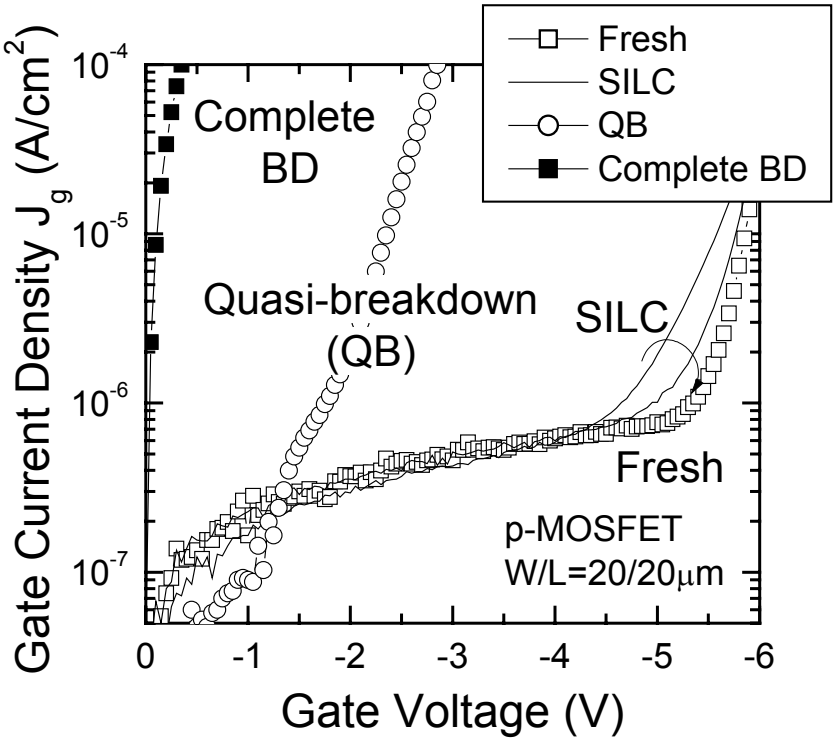


Fig. 2.7 I-V characteristics of gate oxide at various stages of stressing – Fresh, SILC, quasi-breakdown (QB) and complete breakdown.

larger than the stress-induced leakage current but smaller in magnitude than complete a breakdown and post I_g-V_g, which is non-ohmic (Refer to Fig. 2.6 and Fig. 2.7). Leroux *et al.* and Bruyere *et al.* have shown using emission microscopy that QB and complete BD occur at different spatial locations [2.34] and their intrinsic Weibull distributions are different [2.35]. In addition, Pompl *et al.* have also shown that QB and complete BD share completely different temperature and field acceleration behaviors [2.36]. Although the gate oxide has not been totally destroyed, the gate leakage current at the onset of QB is generally far too high for acceptable device operation. Moreover, unlike SILC, the gate voltage fluctuation after the onset of QB, also becomes much more noticeable and noisy, exhibiting both random non-switching 1/*f* noise and multilevel random telegraph switching noise [2.38].

Ever since its discovery in 1994 by Lee and Cho *et al.* [2.10], the conduction mechanism of quasi-breakdown (QB) has been subjected to much controversy and debate. Lee and Cho *et al.* attributed QB to localized physical damage at the anode interface due to energy released by the injected carrier at the anode. QB is triggered when the localized damage region reduces the effective oxide thickness to the direct tunneling regime, allowing carriers to directly tunnel through the oxide [2.10]. Hirose *et al.* attribute the oxide thinning to the formation of localized conduction filament 2-3 nm from the Si/SiO_2 interface [2.30]. On the other hand, Houssa *et al.* and Degraeve *et al.* attribute QB leakage to a percolation path formed due to electron traps linking the anode and cathode [2.43]. The percolation model is able to explain the power-law behavior for the leakage current and temperature dependence of the gate current after QB. It is also able to explain the Weibull slope [2.41] and critical defect density [2.41] for various oxide thicknesses and shows good agreement with the experimental data.
Okada *et al.* attribute QB leakage current to variable range hopping conduction mediated by localized states due to electrical stressing [2.37]. Using this model, the temperature dependence and large fluctuation in current and voltage at QB can be explained. Miranda *et al.,* using a point conduction model, are also able to explain the large leakage current fluctuation as the switching ON/OFF state of one or more local conduction spots [2.46]. Table 2.2 in the next page, shows a brief summary of all the proposed conduction mechanisms for QB.

Table 2.2 Summary of quasi-breakdown conduction mechanism

QB Conduction Model	References
1. Direct-tunneling through locally damaged region	Lee and Cho *et al.* [2.10], [2.26] Y.D. He *et al.* [2.27], [2.28] Hirose *et al.* [2.30], Z. Y. Ting *et al.* [2.31]
2. Percolation through neutral electron traps	Degraeve *et al.* [2.40] J. H. Stathis *et al.* [2.41] Houssa *et al.* [2.43] Vandewalle *et al.* [2.44]
3. Point contact conduction with energy quantum	Miranda *et al.* [2.47] Sune *et al.* [2.48],[2.49]
4. Variable-range hopping (VRH) through localized states	Okada *et al.* [2.37]
5. Analog and digital-mode conduction	Sakura *et al.* [2.38] Tomita *et al.* [2.39]
6. Multiple trap-assisted tunneling (TAT)	Depas *et al.* [2.33]

The key query to the correct model for QB has not been resolved, primarily because of the complex nature of QB and the inability of the various models to explain all the observed phenomena.

2.3.1 Direct Tunneling Model [2.10][2.26]-[2.30]

The direct tunneling model was first proposed by Lee and Cho *et al.* to explain the occurrence of quasi-breakdown (QB) [2.10]. In gate oxides thinner than 50 Å, the traveling distance of electrons in the oxide conduction band after Fowler-Nordheim (F-N) tunneling would be shorter than the electron mean free path (Refer to Fig. 2.8a).

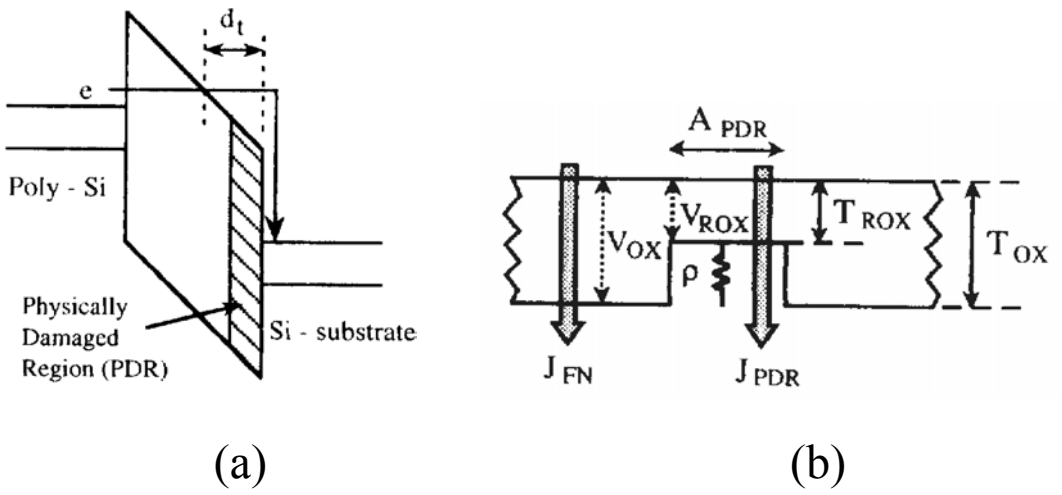


Fig. 2.8 Schematic drawing for (a) electron transport in the ultra thin gate oxide under high field stress (b) Current path in the oxide after quasi-breakdown. (After [2.10])

The electrons after a ballistic transport in the oxide conduction band, release most of its energy at the anode, resulting in a localized physically damaged region (PDR) near the Si/SiO_2 interface. The formation of the PDR reduces the hole barrier and the effective oxide thickness, resulting in a large increase in hot hole injection from the anode. Gate voltage fluctuation at QB is attributed to the dynamic charge trapping and detrapping processes in the PDR. Total gate leakage current $J_{g,}$ is then a superposition of F-N current tunneling through undamaged oxide J_{FN} and the direct tunneling current J_{PDR} through the PDR. They can be described by the following equations:

$$J_g = (1\text{-}A_{ratio}) J_{FN} + A_{ratio} J_{PDR} \tag{2.1}$$

where A_{ratio} is defined as

$$A_{ratio} = A_{PDR} / A_{cap} \tag{2.2}$$

and A_{PDR} is the area of quasi-breakdown spot while A_{cap} is the total capacitor area. At the onset of quasi-breakdown, the total voltage drop in the gate oxide can then be described on the following page in equation (2.3):

$$V_{OX} = V_{ROX} + J_{PDR} \rho (T_{OX} - T_{ROX}) \tag{2.3}$$

where ρ is the finite resistivity of the damaged region and V_{ROX} is the voltage drop across the undamaged region of the QB conduction path as shown in Fig. 2.8(b). Using the direct tunneling model, various groups are able to explain most of the observed phenomena including the observation of a hole current component in n-

MOSFET and p-MOSFET after QB [2.28], multilevel random telegraph signal (RTS) and the annealing behavior of post-QB oxides [2.29].

2.3.2 Percolation Model [2.40]-[2.43]

Another widely accepted model for QB is the percolation model, first proposed by Degraeve *et al.* [2.40] to explain intrinsic oxide breakdown and later applied to ultra-thin oxide soft breakdown phenomenon by Houssa *et al.* [2.43]. It is able to explain both the anode hole injection model (AHI) [2.16] and the electron trap generation model by directly linking the injected holes to the electron trap generation. At quasi-breakdown (QB), the electron traps forms a percolation path linking both the anode and cathode as shown in Fig. 2.9. Conduction between neighboring electron traps becomes possible when the distance between these traps is less than or equal to 0.9 nm [2.40].

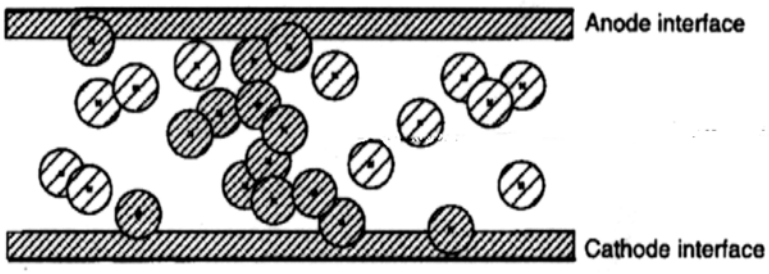


Fig. 2.9 Schematic illustration of percolation model for intrinsic oxide breakdown based on electrons trap generation. Conduction path is indicated by the shaded spheres. (After [2.33])

Assuming that the current between two neighboring traps is proportional to the square of applied voltage, i.e.

$$I = \sigma_{ab}\, v^2 \tag{2.4}$$

where σ_{ab} is the bound conductivity. From this simple quadratic dependence of current and assuming a percolation path, the power–law dependence for the leakage current at QB can be satisfactorily simulated (Refer to Fig. 2.10)

The main strength of the percolation model is its ability to explain the thickness dependence of the Weibull slope. Degraeve *et al.* have shown that at lower

oxide thickness, the percolation model predicts that fewer traps are required to form a breakdown path, thus resulting in a higher likelihood of breakdown [2.40]. The model is also able to fit the slope β of experimental Weibull plot for different oxide thickness, as shown in Fig. 2.11. However, it is important to note that the good fit of the thickness dependent Weibull slope with modeling using the percolation model does not guarantee its validity. In fact, all other models including LPDR and variable range hopping model etc. can also explain the Weibull slope by assuming a different critical defect level for different oxide thickness. Nevertheless, the percolation model provides a simple and relatively accurate modeling of the QB phenomenon without too many implicit assumptions. Using the percolation model, Depas *et al.* [2.33] and Houssa *et al.* [2.43] have shown that the percolation model is able to explain the post-QB I-V characteristics, and its observed temperature and oxide thickness dependence.

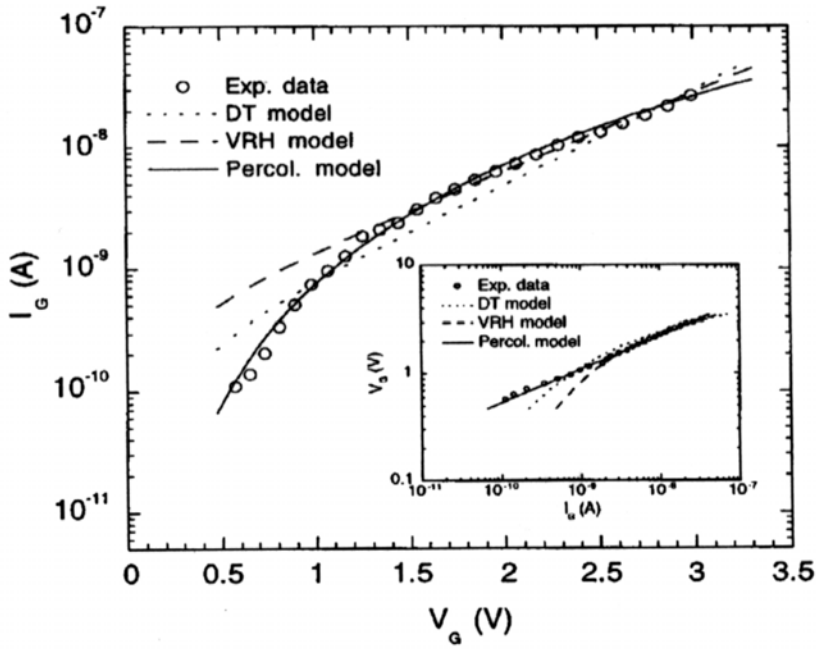


Fig. 2.10 Current-voltage characteristics of a MOS capacitor with a 4.2 nm gate oxide. Solid line is fit obtained with a percolation model. (After [2.43])

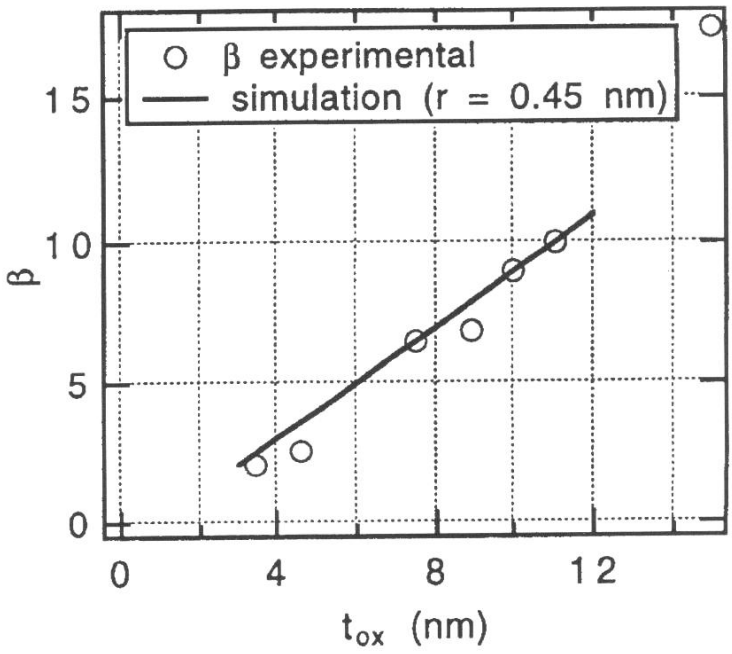


Fig. 2.11 Simulated and measured Weibull slope β for charge-to-breakdown QBD as a function of oxide thickness. (After [2.40])

2.4 Device Scaling and Dielectric Performance

According to the ITRS (International Technological Roadmap Semiconductor) 2003 [2.51], the technological limit for using silicon dioxide (SiO_2) will be reached by 2007 for the 65 nm technology node due to its excessive leakage current. As a thermal oxide, SiO_2 is native to Si with many wonderful attributes. It is thermally and electrically stable on Si, has the ability to form a low defect density interface (~ 10^{10} /cm^2) with Si, has excellent dielectric strength (~10 MV/cm) and a large bandgap of 9 eV [2.52]. These outstanding electrical properties have enabled the microelectronic industry to scale the integrated devices with significant improvements in performance, off-state power consumption, and reliability. At the same time, it also meant that any alternative gate dielectric will face a formidable challenge in replacing SiO_2. In spite of its many superior material characteristics, SiO_2 suffers from a relatively low relative dielectric permittivity of K = 3.9. For a gate dielectric, a low K value means that gate capacitance can only be increased by aggressively thinning down the gate dielectrics as shown in equation (2.5) on the next page.

$$C_{ox} = \frac{K\varepsilon_0 A}{t_{ox}} \tag{2.5}$$

where K, ε_0 is the dielectric constant (also referred to as the relative permittivity of the dielectric to air) and permittivity of air respectively, A is area of the device channel and t_{ox} is the thickness of the dielectrics. Due to device scaling, device area A will be decreased for each new generation and for SiO_2, K is constant at 3.9. Hence the only way gate capacitance can be increased is by reducing the oxide thickness *tox.* A high gate capacitance is required to ensure sufficient drive current in the channel and to overcome short channel effects. As gate length is progressively scaled downwards, short channel effects, including threshold voltage roll-off, become progressively more severe. When SiO_2 thickness is in the region of more than 40 Å, dielectric thickness scaling does not have significant effects on the device characteristics. However, this changes when the gate dioxide becomes very thin (< 30 Å). At an oxide thickness of 11 Å, the gate leakage current is reaching the ITRS limit of 100 A/cm^2 while for low power application, a gate leakage limit of 1 A/cm^2 will be reached for 16 Å SiO_2 . As a result, excessive gate leakage current will threaten the continual scaling of SiO_2 beyond 16 Å. Using theoretical modeling, Tang *et al.* has further shown that the bandgap of SiO_2 begins to decrease when fewer than 3 monolayers of oxide exists [2.53]. This gate SiO_2 thickness limit forces the industry to look at other higher-K material as a potential alternative to SiO_2. Currently, there are many potential material candidates with very high K values, but few, with all the superior characteristics of SiO_2.

2.5 Ultra-thin oxide Reliability

When silicon dioxide is scaled below 15 Å, the oxide degradation mechanism may be changed or certain features become more important. Unlike thicker oxides where occurrence of QB or complete breakdown defines the lifetime of the gate oxides, post-QB in ultra-thin oxides becomes much more relevant and important due to the ability to withstand several QB or degraded spots [2.54]. Fig. 2.12 shows the gate leakage current for gate oxides of different thickness. It is interesting to note that when oxide thickness is thinner than 20 Å, the initial unstressed leakage current is

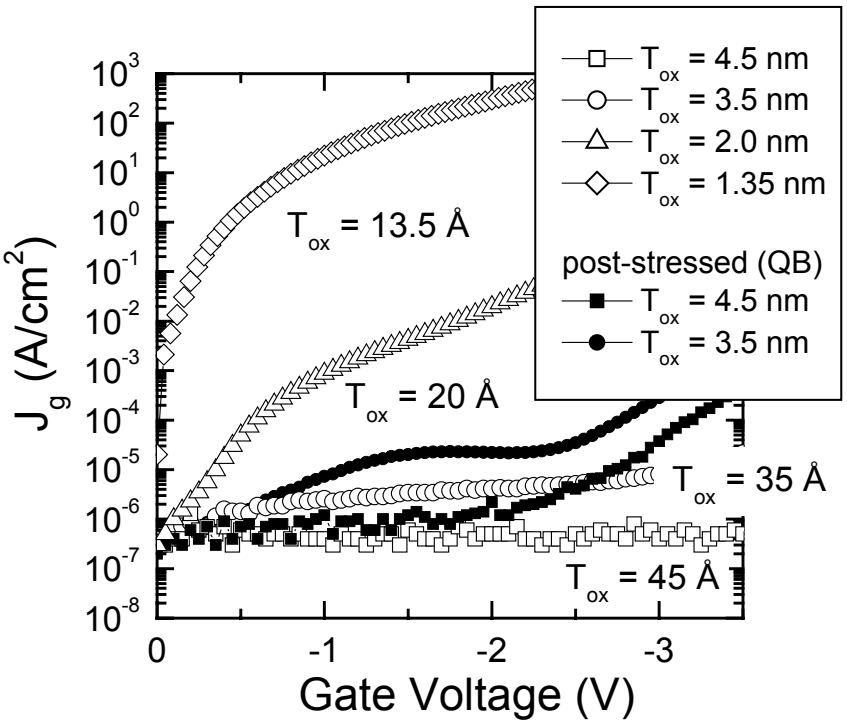


Fig. 2.12 Gate leakage current for different oxide thickness at fresh unstressed state and after post-QB. It can be observed that post-QB leakage current for 35 - 45 Å oxides are actually lower than that of the direct tunneling leakage current of 13 - 20 Å at its initial unstressed state.

higher than that of post-QB leakage in thicker oxides. Due to this high gate leakage for ultra-thin oxides, occurrences of QB may be masked out by the initial gate leakage current and become less critical.

Wu *et al.* was one of the first to observe a gradual change in gate leakage under electrical stress as gate oxide scales below 18 Å and propose gate leakage current density as a criterion for determining oxide lifetime [2.55]. Unlike thicker oxides where discrete increases in gate leakage current were observed at the onset of QB, ultra-thin oxides do not exhibit such discrete current 'jumps'. Instead, a progressive increase in gate leakage is observed before final complete breakdown as shown in Fig. 2.13. Subsequently, many other researchers [2.55]-[2.66] also show similar phenomenon and various statistical models were proposed to explain the wearout and progressive degradation in ultra-thin oxides. Two main issues were raised regarding the wearout modeling of ultra-thin oxides: (1) spatial correlation of the degraded sites and (2) the mechanism of the various stages observed in progressive wearout.

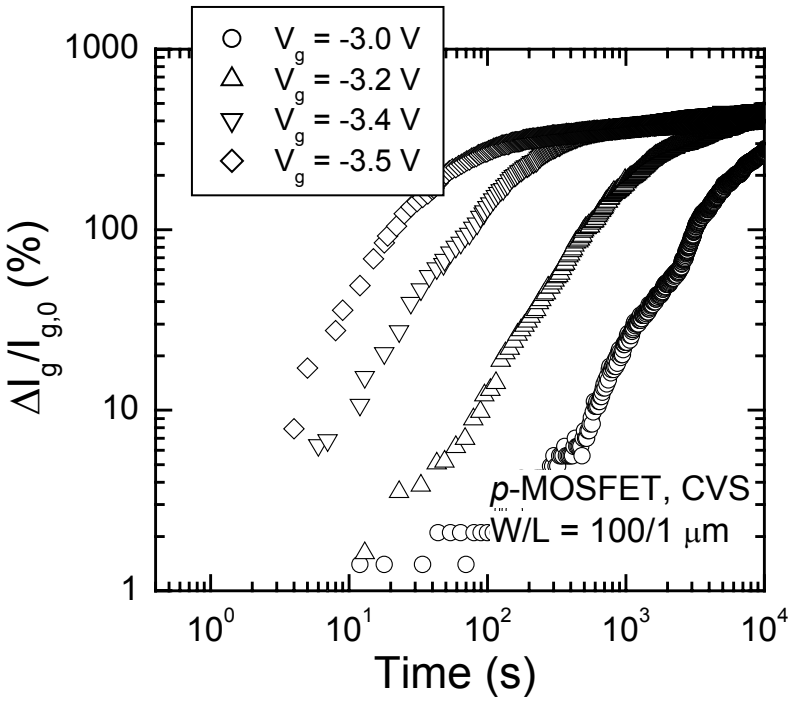


Fig. 2.13 Time evolution of gate current before and after onset of QB for ultra-thin oxides (13.5 Å) under various gate bias CVS.

It can be observed from Fig. 2.13 that there exists two stages of oxide wearout. In the first stage, the gate leakage current increases linearly with stress time on a log-log plot. In the second stage, however, there is a saturation in the increase of gate leakage. For the initial stage, by comparing the gate leakage current after progressive wearout for various gate areas, Mosieur *et al.* observed that the degraded gate leakage current does not increase in tandem with the sample areas. Instead, a highly localized degraded point at the onset of wearout was observed using emission microscopy and this was observed to increase in size over time [2.59]. On the other hand, Alam *et al.* show that for sub-2 nm oxides, the correlation between the initial and subsequent degraded spots is low and the initial progressive wearout can be modeled by multiple degraded spots [2.60]-[2.62].

The mechanism of the initial wearout stage is also not ascertained. It was observed that the wearout is gate voltage and temperature driven. By using substrate bias stressing in the inversion regime, Monsieur *et al.* conclude that wearout is highly substrate bias sensitive and hence is defect driven [2.59]. Farmer *et al.* propose a conduction path with constant degraded characteristics, but with increasing area

[2.64]. On the other hand, Miranda *et al.* have modeled the conduction path using point contact model with increasing area and decreasing barrier height [2.65].

In the second stage, Hosoi *et al.* propose that the saturation occurs due to external resistance effects which eventually limit the gate current increase [2.67]. The verification and ascertainment of these effects will be of great importance to the commercial foundries due to its significant scatter in statistical breakdown for ultra-thin oxide. Moreover, the adoption of a new current density criterion may revolutionize the conventional lifetime prediction methodology and its dependencies on area and gate voltage stress will be the focus of study in chapter 5.

2.6 High-*K* Dielectrics Reliability

Many materials are currently under investigation as potential replacement for SiO_2. Due to high gate leakage current, high-*K* gate dielectrics are needed to reduce gate leakage and still afford reasonable gate capacitance. It has been shown that metal gate (WN_x or TiN) can achieve similar gate leakage as polysilicon with appropriate in-situ nitridation during the sputtering process [2.68]. In addition, metal gate has a superior t_{BD} compared to polysilicon [2.68] and improved boron penetration immunity.

The current reliability studies on high-*K* dielectrics are been mainly focused on high-*K* dielectrics with polysilicon electrode. It has been found that substrate hot hole trapping in p-FETs severely degrades the device. In n-MOSFETs, hot electron trapping occurs only under illumination [2.69]. The experimental results suggest that holes can act as a defect precursor leading to enhanced charge trapping in high-*K*. At the same time, hot carrier reliability studies of n-MOSFETs (HfO_2 with polysilicon gate) indicate longer lifetime when compared to SiO_2 in n-MOSFETs. It is suggested that despite the lower barrier height, HfO_2 gate dielectrics with surface nitridation show more robust interface as evidenced by the lower interface trap generation at the nitrided Si surface [2.71]. Surface nitrided HfO_2 also shows better hot-carrier immunity compared to SiO_2 for the same substrate stressing current [2.72]. Moreover, it has been demonstrated that charge trapping in the high-*K* gate stack is substantial and follows a charge trapping model without creation of additional traps. Dynamic stressing studies on ultra-thin HfO_2 show that dynamic stressing leads to lifetime

enhancement similar to the SiO_2 case. This is attributed to the longer trapping time and detrapping time during the reverse bias [2.73].

2.6.1 High-*K* charge trapping

Charge trapping during electrical stress is observed to be significant for high-*K* dielectrics. Large amount of fixed charge and charge trapping compared to SiO_2 were observed for various types of high-*K* dielectrics deposited using Atomic Layer Chemical Vapor Deposition (ALCVD) [2.69],[2.70], [2.86], Metal Organic Chemical Vapor Deposition (MOCVD) [2.84],[2.85],[2.79], and reactive sputtering using physical vapor deposition (PVD) [2.73],[2.74]. The charge trapping centers responsible for the fixed and mobile charge may be at the interface of the bulk high-*K* and interfacial layer [2.76],[2.77] or within the bulk high-*K* layer itself [2.78], and are highly polarity dependent [2.79]. These charge centers result in hysteresis and significant instability in the threshold voltage, which will pose a serious problem for high-*K* dielectrics implementation. In addition, it has been shown that charge trapping may lead to polarity dependent high-*K* degradation and breakdown [2.78],[2.79].

2.6.2 Stack Reliability

The breakdown mechanism in high-*K* dielectrics has been studied extensively [2.75], [2.80]-[2.86]. It was reported that gate current through high-*K* dielectrics showed both high level and low level fluctuations at the onset of soft breakdown [2.75]. At the same time, previous reports had also indicated that breakdown may occur at either the bulk or interfacial layer of the high-*K* stacks due to its intrinsic multi-layer structure [2.86]. As such, it is believed that the breakdown mechanism of high-*K* stack is much more complicated than single SiO_2 dielectrics.

Bimodal breakdown has been observed since the reliability of high-*K* stack was first attempted. Yamaguchi *et al.* attribute this to a two-stage breakdown process which may have led to dual breakdown modes [2.80]. Using different high temperature annealing, it was observed that serious degradation occurs due to partial crystallization of the high-*K* dielectrics [2.80]. On the other hand, Kauerauf *et al.* [2.81] and Degraeve *et al.* [2.86] have shown that a polarity dependent breakdown exists in ZrO_2/SiO_2 stacks. Using different thickness of interfacial SiO_2 and high-*K* dielectrics, it is observed that the dual polarity dependent breakdown can be attributed

to either a bulk or interfacial breakdown. Lee *et al.* observe that for a thin high-*K* HfO_2/SiO_2 stack, breakdown is limited by interfacial layer degradation instead of the conventional bulk high-*K* layer [2.84] and is polarity independent for HfO_2 with TaN electrode [2.85].

Although the interfacial layer (IL) in a high-*K* stack is unintended, it is also unavoidable due to the high oxygen ambient during high-*K* deposition. The importance of the interfacial layer to the overall high-*K* stack has been further verified by various researchers, through modeling and by experimental reliability testing. Fan *et al.* have concluded the higher dielectric constant IL is required to attain low EOT with reduced leakage current [2.83]. Moreover, the transition region between the high-*K* film and the IL will become more important due to the sacrifice of EOT and leakage current, as this transition region expands at the expense of the high-*K* film. The inconsistency in breakdown mechanism as the high-*K* film is scaled to a thinner dimension can be reconciled by the larger traps generation within the thicker film. Degraeve *et al.* have shown that a high level of intrinsic defect exists in high-*K* stack and this may affect the overall reliability of high-*K* stacks [2.87]. In every case, it is clearly evident that increased charge trapping or trap generation under electrical stressing, leads/ to higher chances of bulk breakdown. The only dissenting evidence against this proposal is found in the study by Zhang *et al.* who found polarity dependent charge trapping and defect generation for thin film with EOT ~ 11 Å [2.79]. In their case, a very high TDDB Weibull slope of 3 is obtained, showing possibility of bulk breakdown in even thinner film. However, since nitridated Hf-silicate is used which is expected to have a higher trap generation, the phenomenon may indicate that nitridation tends to result in higher charge trapping, leading to preferential bulk breakdown even for very thin film.

In retrospect, it can be seen that there are many areas of concern for high-*K* dielectric reliability. In particular, it is observed that the breakdown mechanism as thickness of the high-*K* film is scaled downward is not clearly determined at the moment. At the same time, significant charge trapping is observed for high-*K* dielectrics which have detrimental effects on mobility, threshold voltage stability and possibly, the stack reliability. Several other important issues are also not addressed including the physics of breakdown, impact of plasma damage and process-related

defect optimization for high-*K* dielectrics. These topics are outside the scope of this thesis but are nevertheless very important, and will be proposed for future studies.

2.7 Summary

As gate dielectrics scales downwards, various reliability issues have surfaced. For gate oxides thinner than 50 Å, a phenomenon known as quasi-breakdown was observed prior to complete breakdown. While extensive studies have been conducted, the conduction mechanism of quasi-breakdown remains controversial. Two main models that are widely cited are the direct tunneling and percolation path models. While these and other proposed models are able to explain some of the observed phenomena, a unifying model which can explain all the experimental evidences is solely lacking. In most cases, the validity of a proposed model is based exclusively on the fit with experimental results. Using both experimental evidences and modeling, the full understanding of quasi-breakdown mechanism and its conduction kinetics will be attempted and will be the main focus in this thesis.

For ultra-thin gate oxide below 20 Å, it is shown that oxide degradation even after the onset of QB is relevant and important for device lifetime characterization. It is observed that unlike thicker oxide, QB in ultra-thin oxide (20 Å) may be tolerated and its subsequent degradation can be characterized and modeled using various statistical and physical conduction modeling. Moreover, as the thickness of silicon dioxide reduces even further to 14 Å, QB is not distinctly observed and lifetime prediction using the conventional methodology may be inappropriate. Literature reports have shown that progressive wearout is independent of area and is gate voltage and temperature driven. However, using the new failure criterion based on gate leakage current density, it is to be determined if this is still valid and a comparison with excessively optimistic conventional time-to-complete breakdown needs to be performed.

As gate silicon dioxide outlives its usefulness at around 14 Å, beyond in which direct tunneling leakage current will be too high for general device applications, high-*K* dielectrics is needed for the 65 nm technology node and beyond. Based on current literatures, the mechanism for high-*K* stack breakdown is still uncertain. In comparison to SiO_2, the reliability of high-*K* stack will be very challenging due to the change in conduction mechanism, band structure and intrinsic properties of the high-*K* dielectrics. Despite these difficulties, it is clear that high-*K* dielectrics will be urgently needed and its breakdown mechanism has to be thoroughly investigated for future lifetime prediction and reliability studies.

Reference

[2.1] D. J. Maria, E. Cartier, and D. Arnold, "Impact Ionization, trap creation and breakdown in silicon dioxide films on silicon," *J. Appl. Phys*., vol. 73, p. 3367, 1993.

[2.2] J. Maserjian and N. Zamani, "Observation of positively charges state generation near the Si/SiO_2 interface using Fowler-Nordheim tunneling," *J. Vac. Sci. Technol.,* vol. 20, pp. 743-748, 1982.

[2.3] E. Rosenbaum, and L. F. Register, "Mechanism of stress-induced leakage current in MOS capacitors," *IEEE Trans. Electron Dev*., vol. 43, p. 1499, 1997.

[2.4] D. DiMaria, and E. Cartier, "Mechanism of stress-induced leakage currents in thin silicon dioxide films," *J. Appl. Phys*., p. 3883, 1995.

[2.5] J. Wu, L. F. Register, and E. Rosenbaum, "Trap-assisted tunneling current through ultra-thin oxide," *Int. Reliability Phys. Symp.,* p. 389, 1999.

[2.6] S. I. Takagi, N. Yasuda, and A. Toriumi, "Experimental Evidence of Inelastic Tunneling in Stress-Induced Leakage Current," *IEEE Trans. on Electron Dev*., vol. 46, pp. 335-341, 1999.

[2.7] M. G. Chen, C. H. Liu, M. T. Lee, and K. Y. Fu, "New Experimental Findings on SILC and SBD of Ultra-thin Gate Oxides," *Int. Reliability Workshop*, pp. 114 – 117, 1999.

[2.8] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and G. Ghidini, " Evidence for Recombination at Oxide Defect and New SILC model," *IEEE 38th Annual Int. Reliability Phys. Symp.,* pp. 55-64, 2000.

[2.9] T. Wang, N. K. Zous, J. L. Lai, and C. Huang, "Hot hole stress induced leakage current (SILC) transient in tunnel oxide," *IEEE Electron Dev. Lett*., vol. 19, pp. 411-413, 1998.

[2.10] S. H. Lee, B. J. Cho, J. C. Kim, and S. H. Choi, "Quasi-breakdown of ultrathin gate oxide under high field stress," *IEDM Tech. Dig*., pp. 605-608, 1994.

[2.11] K. F. Schuegraf, D. Park, and C. Hu, "Reliability of thin SiO_2 at direct tunneling voltages," *IEDM Tech. Dig.,* pp. 609-612, 1994.

[2.12] S. Song, J. H. Yi, W. S. Kim, J. S. Lee, K. Fujihara, H. K. Kang, J. T. Moon, and M. Y. Lee, "CMOS Device Scaling Beyond 100nm," *IEDM Tech. Dig.*, pp.235-238, 2000.

[2.13] X. Guo, and T. P. Ma, "Tunneling Leakage Current in Oxynitride: Dependence on Oxygen/Nitrogen Content" *IEEE Electron Dev. Lett*., vol. 19, pp. 207-209, 1998.

[2.14] I. C. Chen, S. E. Holland, and C. Hu, "Electrical breakdown in thin gate and tunneling oxides," *IEEE Trans. Electron Dev.,* vol. 32, p. 413, 1985.

[2.15] K. F. Schuegraf and C. Hu, "Hole injection SiO_2 breakdown model for very low voltage lifetime extrapolation," *IEEE Trans. Electron Dev.,* vol. 41, pp. 761-767, 1994.

[2.16] J. C. Lee, I. C. Chen, and C. Hu, "Modeling and Characterization of Gate Oxide Reliability," *IEEE Trans. Electron Dev*., vol. 35, no. 12, pp. 2268-2277, Dec 1988.

[2.17] D. J. Dumin, J. R. Maddux, R. S. Scott, and R. Subramoniam, "A model relating wearout to breakdown in thin oxides," *IEEE Trans. Electron Dev*., vol. 41, pp. 1570-1580, 1994.

[2.18] J. S. Suehle, "Ultrathin gate oxide reliability : Physical models, statistics, and characterization," *IEEE Trans. Electron Dev.,* vol. 49, pp. 958-969, 2002.

[2.19] J. W. McPherson and D. A. Baglee, "Acceleration factors for thin gate oxide stressing," *Proc. Int. Reliab. Phys. Symp.,* vol. 23, pp. 1-5, 1985.

[2.20] J. W. McPherson and H. C. Mogul, "Underlying physics of the thermochemical E model in describing low-field time-dependent dielectric breakdown in SiO_2 thin film," *J. Appl. Phys*., vol. 84, pp. 1513-1523, 1998.

[2.21] D. J. DiMaria, "Explanation for the polarity dependence of breakdown in ultrathin silicon dioxide films," *Appl. Phys. Lett*., vol. 68, pp. 3004-3006, 1996.

[2.22] D. J. DiMaria, "Defect generation under substrate-hot-electron injection into ultrathin silicon dioxides layers," *J. Appl. Phys.*, vol. 86, pp. 2100-2109, 1999.

[2.23] P. E. Nicollian, W. R. Hunter, and J. C. Hu, "Experimental evidence of voltage driven breakdown models in ultrathin gate oxides," *Proc. Int. Reliab. Phys. Symp*., pp. 7-15, 2000.

[2.24] A. Hiraiwa, S. Sakai, and D. Ishikawa, "Voltage-driven distribution of gate oxide breakdown," *Proc. Int. Reliab. Phys. Symp*., pp. 582-583, 2003.

[2.25] D. J. DiMaria, "Electron energy dependence of metal-oxide-semiconductor degradation," *Appl. Phys. Lett.,* vol. 75, pp. 2427-w428, 1999.

[2.26] B. J. Cho, Z. Xu, H. Guan, and M. F. Li, "Effect of substrate hot-carrier injection on quasibreakdown of ultrathin gate oxide," *J. Appl. Phys*., vol. 86, pp. 6590-6592, 1999.

[2.27] Y. D. He, H. Guan, M. F. Li, and B. J. Cho, "Conduction mechanism under quasibreakdown of ultrathin gate oxide," *Appl. Phys. Lett*., vol. 75, pp. 2432-2434, 1999.

[2.28] H. Guan, M. F. Li, Y. D. He, B. J. Cho, and Z. Dong, "A thorough study of quasi-breakdown phenomenon of thin gate oxide in dual-gate CMOSFETs," *IEEE Trans Electron Dev.*, vol. 47, pp. 1608-1616, 2000.

[2.29] Z. Xu, B. J. Cho, and M. F. Li, "Annealing behaviour of gate oxides leakage current after quasi-breakdown," *Microelectronics Reliability,* vol. 40, p.1341, 2000.

[2.30] M. Hirose, "Electron tunneling through ultrathin SiO_2," *Material Science and Engineering,* vol. B41, pp. 35-38, 1996.

[2.31] David Z. Y. Ting, and T. C. McGill, "Transport analysis of filamentary dielectric breakdown model for metal-oxide-semiconductor tunnel structure," *Int. Workshop Computational Electronics*, pp. 255-258, 1998.

[2.32] M. Depas, T. Nigam, and M. M. Heyns, "Soft breakdown of ultra-thin gate oxide layers," *IEEE Trans. Electron Dev*., vol. 43, pp. 1499-1503, 1996

[2.33] M. Depas, R. Degraeve, T. Nigam, G. Groeseneken, and M. Heyns, "Reliability of Ultra-thin Gate Oxide below 3 nm in the Direct Tunneling Regime," *Jpn. J. Appl. Phys*. vol. 36, pp. 1602-1608, 1997.

[2.34] C. Leroux, D. Blachier, O. Briere, and G. Reimbold, "Light emission microscopy for thin oxide reliability analysis," *Microelectronic Engineering,* vol. 36, pp. 297-300, 1997.

[2.35] S. Bruyere, E. Vincent, and G.Ghibaudo, "Quasi-breakdown in Ultra-thin SiO_2 films: occurrance characterization and reliability assessement methodology," *Proc. 38th In. Reliab. Phys. Symp*., pp. 48-54, 2000.

[2.36] T. Pompl. H. Wurzer, M. Kerber, and L. Eisele, "Investigation of ultra-thin gate oxide reliability behavior by separate characterization of soft breakdown and hard breakdown," *Proc. 38th Int. Reliab. Phys. Symp.*, pp. 40-47, 2000.

[2.37] K. Okada and K. Taniguchi, "Electrical stress-induced variable range hopping conduction in ultrathin silicon dioxides," *Appl. Phys. Lett*., vol. 70, pp. 351-353, 1997.

[2.38] T. Sakura, H. Utsunomiya, Y. Kamakura, and K. Taniguchi, "A detailed study of soft- and pre-soft-breakdown in small geometry MOS structures," *IEDM Tech Dig*., pp. 183-186, 1998.

[2.39] T. Tomita, H. Utsunomiya, T. Sakura, Y. Kanakura, and K. Taniguchi, "A new soft breakdown model for thin thermal SiO_2 film under constant current stress," *IEEE Trans. Electron Dev*., vol. 46, pp. 159-165, 1999.

[2.40] R. Degraeve, G. Groeseneken, R. Bellens, M. Depas, and H. E. Maes, "A consistent model for the thickness dependence of intrinsic breakdown in ultra-thin oxides," *IEDM Tech. Dig.,* pp. 863-866, 1995.

[2.41] J. H. Stathis and D. J. DiMaria, "Reliability projection for ultra-thin oxides at low voltages," *IEDM Tech. Dig*., pp. 167-170, 1998.

[2.42] J. H. Stathis, "Percolation models for gate oxide breakdown," *J. Appl. Phys*., vol. 86, pp. 5757-5766, Nov. 1999.

[2.43] M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, "Soft breakdown in ultrathin gate oxides: correlation with the percolation theory of nonlinear conductors," *Appl. Phys. Lett*., vol. 73, pp. 514-516, 1998.

[2.44] N. Vandewalle, M. Ausloos, M. Houssa, P. W. Mertens, and M. M. Heyns, "Non-Guassian behavior and anticorrelations in ultrathin gate oxides after soft breakdown," *Appl. Phys. Lett.,* vol. 74, pp. 1579-1581, 1999.

[2.45] M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, "Model for the current-voltage characteristics of ultrathin gate oxides after soft breakdown," *J. Appl. Phys*., vol. 84, pp. 4351-4355, 1998.

[2.46] E. Miranada, J. Sune, R. Rodriguez, M. Nafria, X. Aymerich, L. Fonseca, and F. Campabadal, "Soft breakdown conduction in ultrathin (3-5 nm) gate dielectrics," *IEEE Trans. Electron Dev.,* vol. 47, pp. 82-88, 2000.

[2.47] E. Miranda, J. Sune, R. Rodriguez, M. Nafrai, and X. Aymerich, "A Function-Fit model for the soft breakdown failure mode," *IEEE Electron Dev. Lett*., vol. 20, pp. 265-267, 1999.

[2.48] J. Sune and E. Miranda, "Point Contact Conduction at the oxide breakdown of MOS devices," *IEDM Tech. Dig.,* pp. 191-194, 1998.

[2.49] J. Sune and E. Miranda, "Post Soft Breakdown Conduction in SiO_2 Gate Oxides," *IEDM Tech. Dig.*, pp. 533-536, 2000.

[2.50] D. Roy, S. Bruyere, E. Vincent, and G. Ghibaudo, "Electric Field and Temperature Acceleration of Quasi-Breakdown Phenomena in Ultrathin Oxides," *Proc. Int. Reliab. Workshop*, pp. 49-53, 1998.

[2.51] International Technology Roadmap for Semiconductors (ITRS) 2003, Semiconductor Industry Assoc., San Jose, CA., Available : http://public.itrs.net/Files/2003ITRS/Home2003.htm

[2.52] M. L. Green, E. P. Gusev, R. Degraeve, and E. L. Garfunkel, "Ultrathin (< 4nm) SiO_2 and Si-O-N gate dielectriclayers for silicon microelectronics: understanding the processing, structure, and physical and electrical limits," *J. Appl. Phys.,* vol. 90, pp. 2057-2121, 2001.

[2.53] S. Tang, R. M. Wallace, A. Seabaugh, and D. King-Smith, *Appl. Surf. Sci.,* vol. 135, p. 137, 1998.

[2.54] K. Okada, H. Kubo, A. Ishinaga, and K. Yoneda, "A concept of gate oxide lifetime limited by "B-mode" stress induced leakage currents in direct tunneling regime," *Symp. VLSI Technology Dig.,* pp. 57-58, 1999.

[2.55] Y. Wu, D. Bang, G. Lucovsky, and M. R. Lin, "Time Dependent Dielectric Wearout (TDDW) Technique for Reliability of Ultrathin Gate Oxides," *IEEE Electron Dev. Lett.,* vol. 20, no. 6, pp. 262 – 264, 1999.

[2.56] F. Monsieur, E. Vincent, D. Roy, S. Bruyere, J. C. Vildeuil, G. Pananakakis, and G. Ghibaudo, "A Thorough Investigation of Progressive Breakdown in Ultra-thin Oxides. Physical Understanding and Application for Industrial Reliability Assessment," *IEEE Proc. 40th Int. Reliability Phys. Symp. (IRPS)*, pp. 45-54, 2002.

[2.57] E. Farres, M. Nafria, J. Sune, and X. Aymerich, "The statistical distribution of breakdown from multiple breakdown events in one sample," *J. Phys. D*, vol. 24, pp. 407-414, 1991.

[2.58] J. Sune and E. Wu, "Statistics of successive breakdown events for ultra-thin gate oxides," *IEDM Tech Dig.*, pp 147-150, 2002.

[2.59] F. Monsieur, E. Vincent, G. Ribes, V. Huard, S. Bruyere, D. Roy, G. Pananakakis, and G. Ghibaudo, "Evidence for defect-generation-driven wear-out of breakdown conduction path in ultra-thin oxides," *IEEE Proc. 41th Int. Reliability Phys. Symp. (IRPS)*, pp. 424-431, 2003.

[2.60] M. Alam and R. K. Smith, "A Phenomenological theory of correlated multiple soft-breakdown events in ultra-thin gate dielectrics," *IEEE 41th Int. Reliability Phys. Symp. (IRPS),* pp. 406-411, 2003.

[2.61] M. Alam, R. K. Smith, B. Weir, and P. Silverman, "Statistically independent soft-breakdown redefine oxide reliability specifications," *IEDM Tech. Dig.,* pp. 151-154, 2002.

[2.62] M. Alam, R. K. Smith, B. Weir, and P. Silverman, "Uncorrelated breakdown of silicon integrated circuits," *Nature*, 6914, p. 378, 2002.

[2.63] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and M, J. van Duuren, "Correlated defect generation in htin oxides and its impact on Flash reliability," *IEDM Tech. Dig.,* pp. 151-154, 2002.

[2.64] K. R. Farmer, R. Saletti, and R. Buhrman, "Current fluctuations and silicon wear-out in metal-oxide-semiconductor tunnel diode," *Appl. Phys. Lett.*, vol. 52, pp. 1749-1751, 1998.

[2.65] E. Miranda and J. Sune, "Analytic modeling of leakage current through multiple breakdown paths in SiO_2 films," *Proc. Int. Reliab. Phys. Symp.*, pp. 367-379, 2001.

[2.66] B. P. Linder, J. H. Stathis, D. J. Frank, S. Lombardo, and A. Vayshenker, "Growth and Scaling of Oxide Conduction after Breakdown," *IEEE Proc. 41th Int. Reliability Phys. Symp. (IRPS),* pp. 402-405, 2003.

[2.67] T. Hosoi, P. L. Re, Y. Kamakura, and K. Taniguchi, " A New Model of Time Evolution of gate leakage current after soft breakdown in ultra-thin gate oxides," *IEDM Tech. Dig.*, pp. 727-730, 2002.

[2.68] C.W. Yang, Y.K. Fang, S.F. Chen, C.Y. Lin, M.F. Wang, Y.M. Lin, T.H. Hou, L.G. Yao, S.C. Chen, and M.S. Laing, "Effective improvement of high-*K* Hf-silicate/silicon interface with thermal nitridation," *Electronic Lett.*, vol. 39, pp. 421-422, 2003.

[2.69] A. Kumar, T. H. Ning, M. V. Fischetti, and E. Gusev, "Hot-Carrier Charge Trapping and Reliability in high-*K* dielectrics," *Symp. VLSI Technology Dig.*, pp. 152-153, 2002.

[2.70] J. F. Conley Jr., Y. Ono, W. Zhuang, L. Stecker, and G. Stecker, "Electrical properties and reliability of HfO_2 deposited via ALD using $Hf(NO_3)_4$ precursor," *Int. Reliab. Workshop,* pp. 108-112, 2002.

[2.71] P. D. Kirsch, C. S. Kang, J. Lozano, J. C. Lee, and J. G. Ekerdt, "Electrical and spectroscopic comparison of HfO_2/Si interface on nitrided and un-nitrided Si(100)," *J. Appl. Phys.*, vol. 91, pp. 4353-4363, 2002.

[2.72] Q. Lu, H. Takeuchi, R. Lin, T. J. King, C.M. Hu, K. Onishi, R. Choi, C. S. Kang, and J. C. Lee, " Hot Carrier Reliability of n-MOSFET with Ultra-thin HfO_2 Gate Dielectrics and Poly-Si Gate," *Proc. Int. Reliab. Phys. Symp.*, pp. 429-414, 2002.

[2.73] A. Callegari, E. Cartier, M. Gribelyuk, H. F. Okorn-Schmidt, and T. Zabel, "Physical and electrical characterization of Hafnium oxide and Hfanium silicate sputtered films," *J. Appl. Phys*., vol. 90, pp. 6466-6475, 2001.

[2.74] Y. H. Kim, K. Onishi, C. S. Kang, R. Choi, H. J. Cho, S. Krishnan, A. Shahriar, and J. C. Lee, "Dynamic reliability characteristics of ultra-thin HfO_2," *Proc. Int. Reliab. Phys. Symp.,* pp. 46-50, 2003.

[2.75] H. Satake, "Soft breakdown Phenomena in high-*K* gate dielectrics," *Electrochemical Society 204th Meet*, Abstr. 573, 2003.

[2.76] J. C. Wang, S. H. Chiao, C. L. Lee, and T. F. Lei, Y. M. Lin, M. F. Wang, S. C. Chen, C. H. Yu, and M. S. Liang, "A physical model for the hysteresis phenomenon of ultrathin ZrO_2 film," *J. Appl. Phys.,* vol. 92, pp. 3936-3940, 2002.

[2.77] L. Pantisano, E. Cartier, A Kerber, R. Degraeve, M. Lorenzini, M. Rosmeulen, G. Groeseneken, and H. E. Maes, "Dynamics of threshold voltage instability in stacked high-*K* dielectrics : role of interfacial oxide," *Symp. VLSI Technology Dig.,* pp. 163-164, 2003.

[2.78] A. Kerber, E. Cartier, R. Degraeve, P. J. Roussel, L. Pantisano, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, "Charge Trapping and Dielectric Reliability of SiO_2-Al_2O_3 Gate Stack with TiN Electrodes," *IEEE Trans. Electron Dev.,* vol. 50, pp. 1261 - 1269, 2003.

[2.79] J. Zhang, E. Zhao, Q. Xiang, J. Chan, J. Jeon, J. S. Goo, A. Marathe, B. Ogle, M. R. Lin, and K. Taylor, "Polarity dependent reliability of advanced MOSFET using MOCVD nitrided Hf-silicate high-*K* gate dielectrics," *Int. Reliab. Workshop,* p. 92, 2002.

[2.80] T. Yamaguchi, T. Ino, H. Satake, and N. Fukushima, "Novel dielectric breakdown model of Hf-silicate with high temperature annealing," *Proc. Int. Reliab. Phys. Symp.*, pp. 34-40, 2003.

[2.81] T. Kauerauf, R. Degraeve, E. Cartier, B. Govoreanu, P. Blomme, B. Kaczer, L. Pantisano, A. Kerber, and G. Groeseneken, "Toward understanding degradation and breakdown of SiO_2/high-*K* stacks," *IEDM Tech. Dig.,* pp. 521-524, 2002.

[2.82] T. Yamada, M. Moriwaki, Y. Harada, S. Fujii and Koji Eriguchi, "The metal gate MOS reliability with improved sputtering process for gate electrode", *IEDM Tech. Dig.*, pp. 319 – 322, 1999.

[2.83] Y. Y. Fan, Q. Xiang, J. An, L. F. Register, and S. K. Banerjee, "Impact of interfacial layer and transition region on gate current performance for high-*K* gate dielectric stack : Its tradeoff with gate capacitance," *IEEE Trans. On Electron Dev.*, vol. 50, pp. 433 – 439, 2003.

[2.84] S. J. Lee, H. F. Luan, C. H. Lee, T. S. Jeon, W. P. Bai, Y. Senzaki, D. Roberts, and D. L. Kwong, "Performance and reliability of ultra-thin CVD HfO2 Gate dielectrics with dual Poly-Si Gate electrode," *Symp. VLSI Technology Dig.,* p. 133, 2001.

[2.85] S. J. Lee, S. J. Rhee, R. Clark, and D. L. Kwong, "Reliability projection and polarity dependence of TDDB for ultra thin CVD HfO2 gate dielectrics," *Symp. VLSI Technology Dig.,* p. 78, 2002.

[2.86] R. Degraeve, T. Kauerauf, A. Kerber, E. Cartier, B. Govoreanu, Ph. Roussel, L. Pantisano, P. Blomme, B. Kaczer, and G. Groeseneken, "Stress polarity dependence of degration and breakdown of SiO_2/high-*K* stack," 41st *Proc. Int. Reliab. Phys. Symp*, pp. 23-28, 2003.

[2.87] R. Degraeve, A. Kerber, Ph. Roussel, E. Cartier, T. Kauerauf, L. Pantisano, and G. Groeseneken, "Effect of bulk trap density on HfO2 reliability and yield," *IEDM Tech. Dig*., pp. 935-938, 2003.

Chapter 3

Measurement Setup and Techniques

3.1 Measurement Techniques

The measurement methodologies used in this thesis consists of two main measurement techniques. Direct-current current-voltage (DCIV) measurement technique [3.2], first proposed by Neugroschel and Sah *et al*., was used predominantly to monitor the interface traps and to a lesser extent, the bulk traps generated in the oxide. Unlike the conventional charge pumping (CP) measurement technique, DCIV uses purely direct-current (dc) measurement, resulting in easy measurement and high detection sensitivity. This technique is used throughout Chapter 4, 5, and 6 to separately monitor the generation of interface traps and oxide bulk traps under Fowler-Nordheim (FN) stressing. Another technique used in this thesis is carrier separation for both n- and p-MOSFET under channel inversion, to separately monitor the tunneling components of electron and hole through the gate dielectrics. This technique is used in a novel way in Chapter 7, to differentiate the interfacial layer and bulk breakdown in high-K dielectrics. Both techniques form the cornerstone of this thesis and will be described in greater detail in the following sections.

3.1.1 Direct Current Current-Voltage (DCIV) Technique

The principle of Direct Current Current-Voltage (DCIV) method on CMOSFETs was described by Neugroschel *et al.* [3.2] and later used in various applications by various researchers. These include DCIV as a nondestructive monitor for implant-induced damage in deep submicron MOSFETs [3.3], extractor for impurity concentration in various regions of MOSFETs [3.4], process diagnosis [3.5], and as a primary monitor of interface and oxide trap density in various oxide reliability studies [3.6]-[3.9]. For DCIV measurement, interface and oxide traps were monitored using the vertical parasitic p-n-p BJT structure present in the p-MOSFET. The technique utilizes a channel controlled by BJT structures which is present in CMOS circuitry. Fig. 3.1 shows a cross-sectional view of a p-MOSFET using surface

control or lateral BJTs in the DCIV measurement configuration. When the lateral BJT is in the forward bias mode with p+source as the emitter, n-well as the base and p-substrate as the collector under the following bias conditions: V_{BE} = 0.3 to 0.6 V, V_C = 0V, V_B = 0 V, the base current measures the recombination current via the interface traps generated during processing or subsequent electrical stress when gate bias is swept from channel accumulation to depletion. This base current I_B, includes three recombination current components: (1) I_{B1} due to defect within the base and emitter – base interface, (2) I_{B2} due to interface traps at the channel and drain region of the Si-SiO_2 interface and (3) I_{B3} due to recombination or tunneling current from interface traps under sidewall spacer. Due to their spatial differences, each component reaches a maxima at different gate bias. At close to zero gate voltage, recombination at interface traps within the channel region, (2) is dominant while (1) and (3) is almost negligible when compared to (2) [3.6].

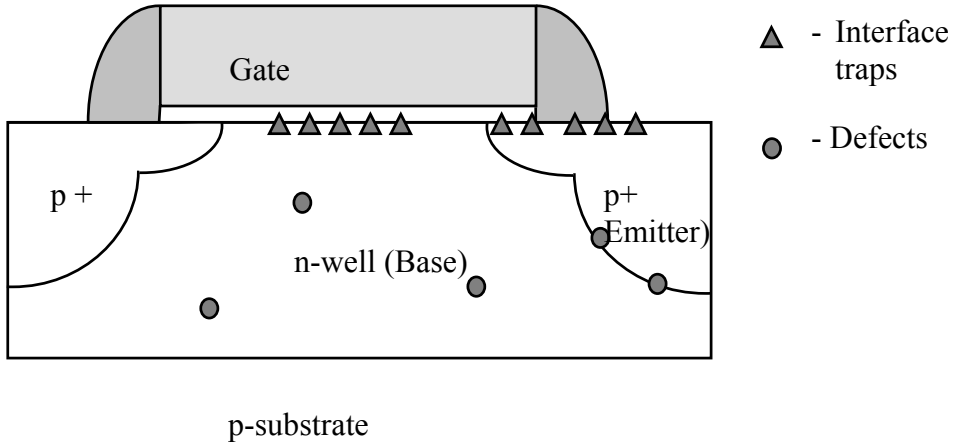


Fig. 3.1 p-MOSFET in top emitter-base configuration with spatial distribution of interface traps and recombination traps centers as shown.

When gate voltage is swept from channel accumulation to depletion, the recombination region shifts from source/drain and extends towards the channel region. The recombination of the injected excess minority carrier with the majority carrier at the interfacial traps at the channel region and at source/drain overlap is thus modulated by the gate bias and reaches a maxima close to intrinsic midgap when $N_s c_e = P_s c_h$ (where c_e and c_h are electron and hole capture rate coefficient, N_s and P_s are the steady state electron and hole steady-state concentration at the SiO_2/Si interface). By applying Shockley-Read Hall statistics, it is possible to obtain the steady state recombination current due to interface traps at both the channel and drain region. I_B is

then related to the number of interface traps N_{IT} and the recombination rate of an interface trap. Assuming that the majority of effective interface traps at the intrinsic Fermi level were at midgap and that both electron and hole capture coefficient is the same as denoted by c_e^t, it can be shown that the base current for the channel I_B is related to the interface trap N_{IT} by the following equations [3.2]:

$$I_B = \frac{qc_e^t n_i^2 \cdot N_{IT}\left[\exp(-\frac{qV_{BE}}{kT}) - 1\right]}{p_{no}\exp(\frac{-q(V_{BE}+V_s)}{kT}) + n_{no}\exp(\frac{qV_s}{kT}) + 2n_i} \tag{3.1}$$

The base recombination current I_B has a peak at

$$I_{B,\max} = \frac{qn_i N_{IT}}{2}\left(\exp\left(-\frac{qV_{BE}}{kT}\right)+1\right) \tag{3.2}$$

and corresponding gate voltage at peak I_B

$$V_{GB} = V_{FB} + V_{S,\max} + \varepsilon_S E_{S,\max}/C_{OX} - \frac{Q_{OT}+Q_{IT}}{C_{OX}} \tag{3.3}$$

where V_{GB} is the surface potential at peak $I_{B,max}$, $E_{S,max}$ is the corresponding surface electric field and C_{ox} is the oxide capacitance per unit area. It can be seen from Eq. (2.8) that the shift of V_{GB} , ΔV_{GB}, is then due to the change in Q_{OT} and Q_{IT}. However, from Shockley-Read-Hall (SRH) statistics, base recombination current will be maximized close to mid-gap level (where electron and hole concentrations are approximately equal) for most effective generation/recombination centers [3.12]. At this mid-gap level, with peak $I_{B,max}$, net charge due to Q_{IT} is nearly zero since it is recognized that the/for interface traps above the intrinsic Fermi level $E_{i,}$ are acceptors and those below are donor type [3.13] and these compensate each other at mid-gap. The peak recombination base current $I_{B,max,}$ then shows the interface traps Q_{IT} while the shift in V_{GB}, ΔV_{GB}, shows purely the oxide trap charge Q_{OT}.

Fig. 3.2 shows the DCIV measurement setup for a four contacts lateral p-channel MOS transistor. Its source is forward biased at 0.3 V while n-well and p-substrate are grounded. A gate voltage sweep from channel accumulation to depletion is applied and base recombination current measured via n-well contact is monitored. This base recombination current I_B measures the recombination of minority holes and

majority electrons along the channel interface and is proportional to the density of interfacial traps at the SiO_2-Si interface.

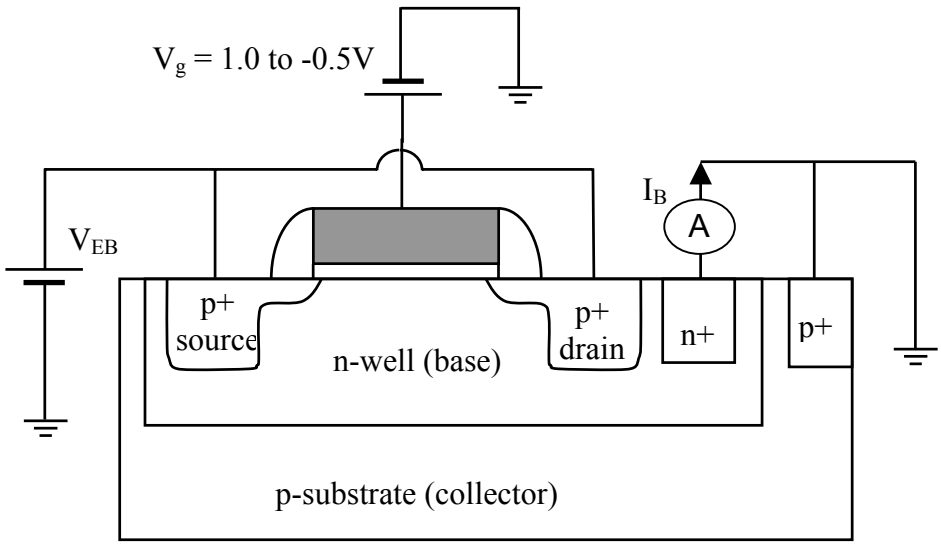


Fig. 3.2 Basic experimental setup for DCIV measurement using p-MOSFET. A vertical parasitic p/n/p-BJT is used with p+ source as emitter, n-well as base, and p-substrate as collector. Forward bias condition of V_{EB} = + 0.3 V and V_{BC} = 0 V is applied and base recombination current I_B is monitored as a function of gate voltage V_g sweep. Drain can be connected together with source as shown or floated.

Unlike the traditional C-V technique, the DCIV technique is able to determine the interface trap generation even in very small transistors and is especially useful in the monitoring of localized phenomenon like quasi- and hard breakdown. It is also able to measure non-uniform interface and oxide traps generated due to hot carrier stressing.

3.1.2 Charge Pumping (CP) Measurement Technique [3.14]

Charge pumping technique was first proposed by Brugler *et al.* [3.14] and has been used in various applications, mainly in plasma process characterization and process optimization, similar to that of DCIV. It is mainly used for the profiling of interface traps and oxide traps along the channel for both process and reliability studies. It has high sensitivity and applicable for small dimension devices with good spatial capability. Fig. 3.3 shows the experimental setup of the CP measurement. AC signal is applied to the gate to drive the channel into inversion and accumulation at different frequency. Under channel inversion, minority electrons from source/drain

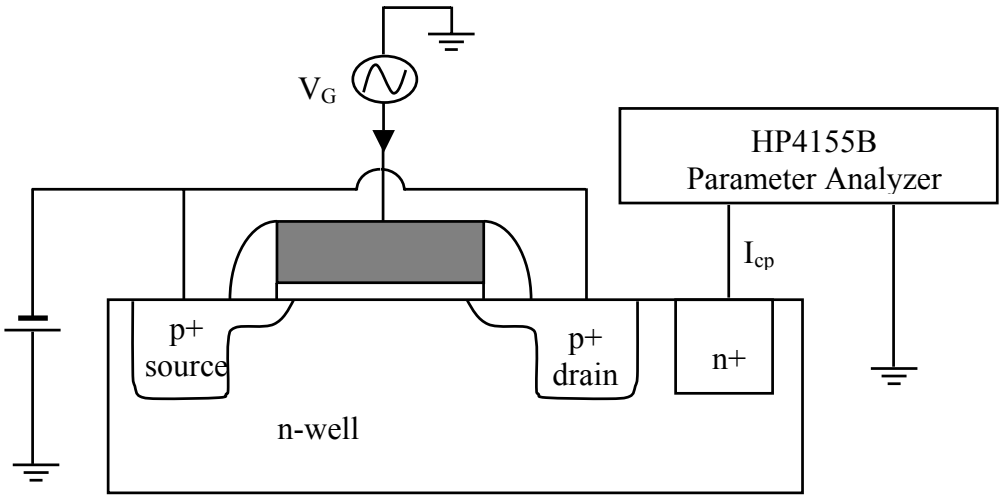


Fig. 3.3 Basic experimental setup for conventional charge-pumping measurements on p-MOSFETs. Source and drain are shorted and lightly reversed biased. An AC signal is applied to gate to alternately drive electrons and holes into the interface traps located at the SiO_2-Si interface. During channel accumulation, recombination of free carrier with the trapped charges cause a net DC substrate current – charge pumping current I_{cp} which is proportional to the areal interface trap density.

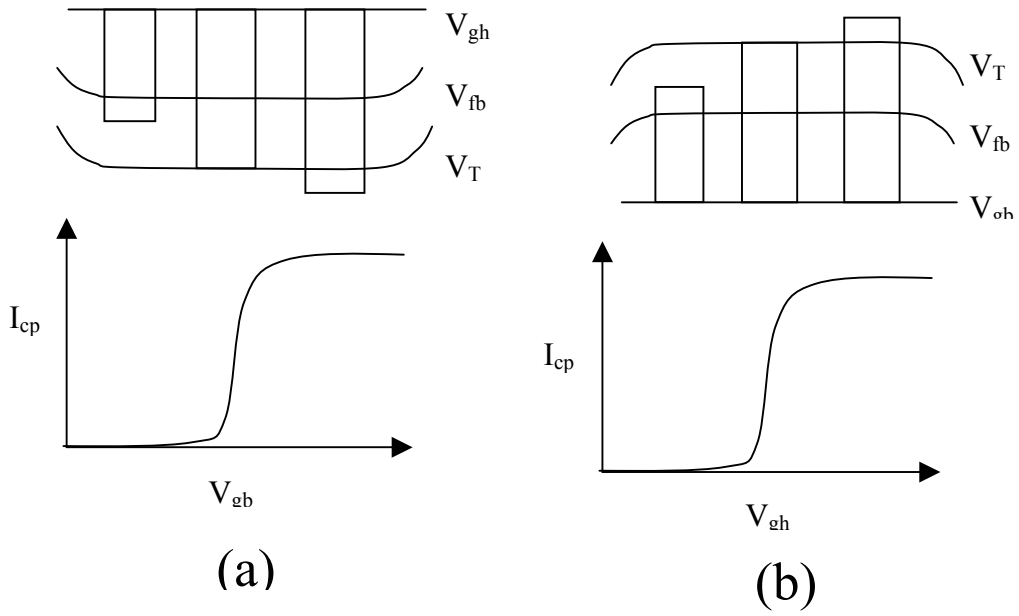


Fig. 3.4 Schematic illustration of CP technique applied to p-channel MOSFETs. (a) Fixed top level V_{gh} with variable V_{gb} and (b) fixed base level V_{gb} and variable V_{gh}. Both configuration shown with the associated charge pumping current versus the variable gate voltage as shown in the lower figures.

were injected into the channel and are trapped by the interfacial traps. The excessive minority carriers are then caused to flow back to the drain/source during the next cycle of channel accumulation. Recombination of the trapped minority and majority carriers under accumulation condition occurs, resulting in a net DC substrate current I_{cp}, due to the alternating charge pumping. A theoretical model of charge trapping technique has been presented by Groeseneken *et al*. using variable base level V_{gb} but with a fixed top level V_{gh}. The fixed V_{gh} is higher than flatband voltage so that charged pumping current I_{cp} flows only when the variable gate voltage goes below the threshold voltage as shown in Fig. 3.4(a). Alternatively, a fixed base level V_{gb} with variable V_{gh} can be used and CP current flows when variable gate voltage is above threshold as shown in Fig. 3.4(b). It can be shown that charge pumping current I_{cp} can then be expressed as:

$$\Delta I_{cp}(X) = qfW \int_{X_{ss}}^{X_d} \Delta N_{IT}(x) \partial x \tag{3.4}$$

where f is the frequency of the applied gate voltage, W is the channel width, and $\Delta N_{IT}(x)$ is the incremental interface trap distribution at position x in the channel. The charge pumping current I_{cp} saturates as the variable amplitude of the gate pulse exceed threshold voltage and this current $I_{cp,max}$ is then directly proportional to the interface trap distributed throughout the gate length as shown by the above equation (3.4).

Many variations of CP technique exist, including varying pulse-top, varying pulse-base etc. However, most suffer from the transient effects inherent in this technique and a separate charge neutralization stage is necessary. At the same time, as gate oxide thickness is scaled downward, gate leakage increases and will mask out or distort the charge pumping signal, necessitating either numerical or measurements compensation [3.15]. In spite of these issues, CP technique has been widely used and accepted as a sensitive and accurate measurement for interface traps.

3.1.3 Comparison between DCIV and CP Technique

The sensitivity of DCIV and reliability of this technique has been the subject of several preceding papers. Neugroschel *et al.* have shown that the sensitivity of DCIV spectra is governed by the base recombination current measurement and in a typical top emitter configuration (whereby the drain/source acts as emitter, the well as base and the substrate as collector) for a p-MOSFET, its minimum I_B detected is better

than 1 pA, translating to an interface trap density $N_{IT} \leq 10^9$ cm^{-2} [3.2]. This is order of magnitude is better than a typical capacitance-voltage (C-V) measurement technique and comparable to that of conventional charge pumping technique.

Jie *et al.* [3.8] and Goh *et al*. [3.9] have both separately compared the measurement results of both charge pumping (CP) measurement and DCIV measurements when applied to Fowler Nordheim stressing and have found a linear relationship between the two methodologies, showing good correlation between both methods in characterizing interface trap generation under uniform electrical stress. Due to the alternating current component in CP measurement, a geometric effect is observed especially in a larger device [3.10]. This originates from the presence of excess minority carrier when gate bias is pulsed from inversion to accumulation, some of which do not have sufficient time to flow to source/drain before the arrival of the majority carrier. These recombine with that of the majority carrier resulting in an excess CP component. In contrast, DCIV which is a purely direct-current, do not suffer from such transient effects. In addition, DCIV can easily distinguish the various interfaces at channel, source/drain overlap and bulk recombination centers in source/drain, through either gate or drain sweep bias. Nevertheless, extraction of interface trap density $N_{IT,}$ using DCIV may not be straightforward due to the following factors:

1. Spatial lateral non-uniformity of bulk recombination centers and
2. Uncertainty of electron and hole capture cross-section.
3. High gate leakage which will add an additional gate leakage component to the base recombination current $I_{B.}$

Current robust fabrication processes have greatly reduced charge lateral non-uniformity with significant improvement to interfacial SiO_2-Si properties. However, device scaling has also resulted in significantly thinner gate oxide. When oxide thickness is reduced beyond 20 Å, direct tunneling from gate to well may supersede the recombination current, hence masking out the recombination current component. However, this can be solved by increasing the emitter forward bias (so as to increase the recombination signal current) or by subtracting the gate leakage current for the actual recombination current as demonstrated by Chung *et al*. [3.11]. Using an optimized forward bias, interface trap measurements using DCIV technique, have

been demonstrated on complementary MOSFETs with a gate silicon dioxide thickness reduced to 13.5 Å.

3.1.4 Carrier Separation

The carrier separation measurement technique has been used in a variety of applications for determining the impact ionization quantum yield γ, extraction of oxide thickness and energy relaxation in inelastic tunneling [3.16] to analyze the conduction mechanism of stress induced leakage current (SILC) [3.17] and quasi-breakdown (QB) [3.1]. In this thesis, the carrier separation measurement technique using the HP4155B semiconductor parameter analyzer was applied to differentiate the electron and hole current component under both substrate inversion and accumulation conditions.

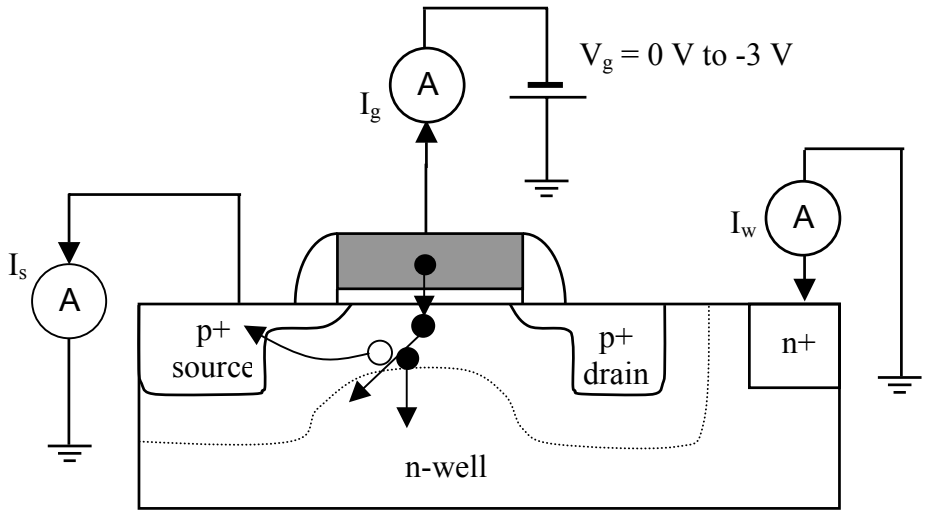


Fig. 3.5 Basic experimental setup for carrier separation measurement on p-channel MOSFETs under inversion mode. Drain is floated while source and n-well are grounded. Gate voltage is swept from 0 V to negative 3 V.

Fig. 3.5 shows the experimental setup for carrier separation on p-channel MOSFET. Carrier separation measurement for a n-channel MOSFET is similar but with reverse bias polarity to the setup as shown above. In this thesis, both n-MOSFETs and p-MOSFETs were used. To prevent the channel from forming between source and drain, the drain contact is left floated while source and substrate are all grounded. Under substrate inversion conditions, ($V_g > 0$ V for n-MOSFET and $V_g < 0$ V for p-MOSFETs), the device acts as an effective p-MOSFET (n-MOSFET) and the source contact measures the hole (electron) while the substrate contact measures the electron (hole) current component for p-MOSFETs (n-MOSFETs). The principles for

gate leakage in conventional silicon dioxides (37 Å) under both inversion and accumulation had been thoroughly examined by Guan *et al*. [3.1]. Fig. 3.6 shows the band diagram of p- and n-channel MOSFETs under carrier separation measurment in the inversion mode.

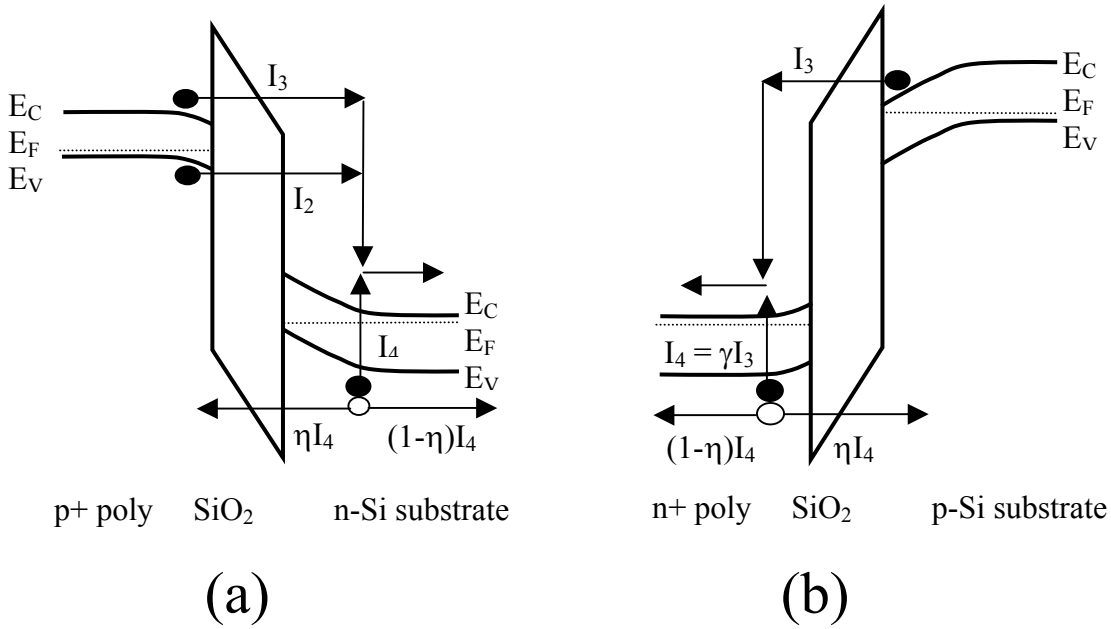


Fig. 3.6 Schematic band diagram (a) of p-channel MOSFET in inversion mode and (b) n-channel MOSFET measured in inversion mode. [After [3.1]]

Based on carrier separation measurements on fresh device, Guan *et al.* have shown that gate electron current I_3 and I_2 flows to the substrate and produces an electron hole pair with unity quantum yield γ=1 [3.1], [3.18]. In this case, $I_{s/d}$ (which is shown by $(1-\eta)I_4$) $\approx I_g$ (represented by I_3, I_2, and ηI_4) and I_w (which is the sum of I_3, I_2, and I_4) is equivalent to $|I_g + I_{s/d}|$ where $\eta \approx 0$. In the case of n-channel MOSFETs under inversion mode, source/drain current $I_{s/d}$ (shown by I_3) is almost equivalent to gate current I_g (I_3+I_4), showing that I_4 is almost negligible.

While carrier separation is only valid under inversion mode, electron and holes leakage path under accumulation is also considered in this thesis with certain assumptions made. In accumulation condition, due to the recombination current between the substrate and gate current, it is difficult to distinguish the separate component of electrons and holes using the carrier separation technique. However, we can consider the gate, substrate and source as a quasi-bipolar transistor with unity minority carrier base transport factor. In this case, for p-channel MOSFETs, the source

acts as a collector of minority holes (electrons) whereas the substrate acts as the base measuring the majority electron current with unity recombination of electron-hole pair in the substrate. This will be used uniquely in Chapter 7 for a novel method of differentiating bulk and interfacial layer breakdown. Moreover, by using metal gate coupled with high-K dielectrics in Chapter 7, it is noteworthy to mention that hole lifetime is significantly reduced in this case, thus eliminating any hole injection from the gate under positive gate bias. As a result only conduction band and valence band electrons are injected into the gate dielectric for p-MOSFET under positive gate bias.

3.2 Summary

Direct current current-voltage (DCIV) measurement technique used for monitoring interface traps and oxide bulk traps, is introduced in this chapter. Using lateral bipolar junction transistor structure present in conventional MOSFETs, DCIV measurement technique employs a simple four point measurement which can accurately measure the interfacial traps and oxide bulk traps in the gate dielectrics, especially those with energy level close to mid-gap. Compared to conventional methods like charge pumping technique, DCIV measurement requires only DC measurement and hence affords ease of implementation with reasonable sensitivity and accuracy.

Another commonly used technique for oxide reliability study: carrier separation technique is also introduced. Using depletion region formed between the channel and p- or n-well, the carrier injected into the gate can be separated into its electron and hole components. This allows us to separately determine the role of both electron and hole in breakdown studies, which is the dominant or key component triggering the initial breakdown.

Both techniques will be used extensively in the subsequent chapters and will provide a very useful insight into the breakdown mechanisms of conventional oxides, as well as high-*K* dielectrics.

Reference

[3.1] H. Guan, M. F. Li, Y. D. He, B. J. Cho, and Z. Dong, "A thorough study of quasi-breakdown phenomenon of thin gate oxide in dual-gate CMOSFETs," *IEEE Trans Electron Dev.*, vol. 47, pp. 1608-1616, 2000.

[3.2] A. Neugroschel, C. T. Sah, M. Han, M. S. Carroll, T. Nishida, J. T. Kavalieros, and Y. Lu, "Direct-current Measurements of Oxide and Interface Traps on Oxidized Silicon," *IEEE Trans on Electron Dev*., vol. 42, pp. 1657-1662, 1995.

[3.3] H. Liao, D. S. Ang, and C. H. Ling, "A comprehensive study of Indium implantation-induced damage in deep submicrometer nMOSFET: Device characterization and damage assessment," *IEEE Trans on Electron Dev.,* vol. 49, pp. 2254-2261, 2002.

[3.4] Y. Wang, and C. T. Sah, "Extraction of impurity concentration profiles by DCIV method," pp. 29-32, 2001.

[3.5] C. T. Sah, "DCIV diagnosis for submicron MOS Transistor design, process, reliability and manufacturing," pp.1-15, 2001.

[3.6] B. B. Jie, M. F. Li, C. L. Lou, W. K. Chim, D. S. H. Chan, and K. F. Lo, " Investigation of interface traps in LDD p-MOSFETs by DCIV method," *IEEE Electron Dev. Lett*., vol. 18, pp. 583-585, 1997.

[3.7] H. Guan, B. J. Cho, M. F. Li, Y. D. He, Z. Xu, and Z. Dong, "A study of Quasi-breakdown mechanism in Ultra-thin gate oxide using DCIV Technique," *Proc. 7th Int. Phys. Failure Analysis*, pp. 81-84, 1999.

[3.8] B. B. Jie, K. H. Ng, M. F. Li, and K. F. Lo, "Correlation between charge pumping method and direct-current current voltage method in p-type metal-oxide semiconductor field-effect transistors," *Jpn. J. Appl. Phys.,* vol. 38, pp. 4696-4698, 1999.

[3.9] Y. H. Goh, and C. H. Ling, "Hot carrier degradation study in PMOSFET using gated-diode drain current and charge pumping measurements," *Proc. ICSE*, Nov. 1996.

[3.10] M. Gaitan, E. W. Enlow, and T. J. Russell, "Accuracy of the charge pumping technique for small geometry MOSFETs," IEEE Trans. Nucl. Sci., vol. 36, pp. 1990-1997, 1989.

[3.11] Steve S. Chung, S. J. Chen, and D. K. Lo, "Interface characterization methodology for nano-CMOS reliability process and device reliability monitors," *10th Proc. Int. Phys. Failure Analysis,* pp. 127-130, 2003.

[3.12] S. M. Sze, *Physics of Semiconductor Devices*, 2nd edition, 1981.

[3.13] G. A. Scoggan, and T. P. Ma, "Effects of electron-beam radiation on MOS structures as influenced by the silicon dopant," *J. Appl. Phys.,* vol. 84, p. 294, 1977.

[3.14] J. S. Brugler, and P. G. A. Jespers, "Charge pumping in MOS devices," *IEEE Trans. Electron Dev*., vol. 16, p. 297, 1969.

[3.15] P. Masson, J-L Autran, and J. Brini, "On the tunneling component of charge pumping current in ultrathin gate oxide MOSFET's," *IEEE Electron Dev. Lett.,* vol. 20, pp. 92-94, 1999.

[3.16] A. Shanware, J. P. Shiely, and H. Z. Massoud, " Extraction of Gate Oxide thickness of n- and p-channel MOSFETs below 20Å from the substrate current resulting from valence-band electron tunneling," *IEDM Tech. Dig.,* pp. 815-818, 1999.

[3.17] S. I. Takagi, N. Yasuda, and A. Toriumi, "Experimental Evidence of inelastic Tunneling and new I-V model for stress-induced leakage current," *IEDM Tech. Dig.,* pp. 323-326, 1996.

[3.18] C. Chang, C. Hu, and R. W. Brodersen, "Quantum yield of electron impact ionization in silicon," *J. Appl. Phys*., vol. 57, pp. 302-309, 1985.

Chapter 4

Bipolar Stressing, Trap Generations and Quasi-breakdown Mechanism Study

4.1 Introduction

For the past decades, continual gate oxide scaling due to device dimension scaling has resulted in different oxide degradations and breakdown mechanisms that are not observed in the thicker oxides. As shown in Chapter 2, quasi-breakdown (QB) [4.1], also named soft breakdown (SBD) [4.2] or B-mode SILC [4.3], has become an important reliability concern due to its increased prevalence as gate oxide thickness scales downwards. Consequently, its breakdown and degradation mechanisms have also been the focus of numerous research studies. Among the various conduction mechanisms proposed for QB are – enhanced direct tunneling due to physical damage region (PDR) [4.1], percolation path model [4.2], multiple TAT, variable range hopping (VRH) [4.3], conduction filament, and point contact conduction [4.4]. Numerous studies have been performed to determine the types of traps triggering QB but have shown conflicting results so far. Guan *et al*. have reported a constant level of interface traps at QB under various stress conditions [4.5] which seems to support the PDR model. On the other hand, Halimaoui *et al.* have observed positive trapping with a QB leakage current that is oxide thickness independent [4.6]. Recently, Sune *et al*. have attempted to simulate the localization of QB spots using the Quantum Point Contact model, obtaining very good fit using a modified percolation path model [4.4]. It is obvious that the key to understanding the QB mechanism lies in the understanding of the degradation mechanism leading to QB. In this aspect, it would be crucial to determine whether if interface traps or oxide bulk traps lead to QB. It is also worthy to note that both complete breakdown and stress-induced leakage current (SILC) are correlated to neutral oxide trap creation. Studies by Sune *et al*. suggest that QB and complete breakdown (CB) have the same common physical origin with similar breakdown statistical distribution [4.7] although this has been disputed by

other researchers who found that different types of traps are involved for the two phenomena. Based on post-QB leakage current and the degree of joule heating, Tomita *et al*. have further characterized QB into two different types - digital and analog [4.4],[4.8].

This chapter deals with two important aspects of thin gate oxide reliability – quasi-breakdown and the effects of bipolar stressing, and is organized as follows: Section 4.1 summarizes the most current findings for quasi-breakdown and the dynamics of bipolar stressing. Section 4.2 briefly elaborates on the experimental setup for the direct-current current-voltage (DCIV) measurement technique and the carrier separation as well as the bipolar electrical stressing used in this experiments. Section 4.3 studies the various stages observed in QB while sections 4.4 – 4.6 examine the trap generation prior to QB and in post-QB oxides. Section 4.7 describes the effect of bipolar stressing on QB and its impact on lifetime prediction. Bipolar current stressing was examined and its effect contrasted with unipolar current stressing. The results indicate that bipolar current stressing has effects not entirely similar to the summation of negative and positive current pulses. It was also observed for the first time that quasi-breakdown (QB) can be characterized into two stages – recoverable and unrecoverable QB. In recoverable QB, electrical recovery is continuously observed while in unrecoverable QB, gate voltage (constant current stress) or leakage current (constant voltage stress) becomes very stable without any electrical recovery observed. Bipolar current stressing also shows that charge-to-quasi-breakdown (Q_{QB}) is distinctly different using bipolar stressing and unipolar stressing on small area samples but not on larger area samples. Sections 4.8 – 4.9 thoroughly investigate the tunneling mechanism of post-QB oxides using carrier separation. In this aspect, Guan *et al*. have presented a thorough study of the conduction mechanism in post-QB oxides using carrier separation measurement. By using a similar methodology but on different stages in post-QB degradation, it was observed that the conduction mechanism in oxide at QB can be explained using direct tunneling of both electrons and holes [4.9]. Finally, section 4.11 summarizes the findings observed in this study and lays down a hole induced QB model which can explains *all* the experimental results observed so far.

4.1.1 Bipolar and Unipolar Current Stressing

While much effort has been concentrated on static stressing, current emphasis has been on dynamic or bipolar stressing. This is because bipolar stressing is more akin to the actual operating condition especially for devices such as EEPROM. The changes in polarity in bipolar stressing have introduced various phenomena which are not observed in unipolar or static stressing. In particular, it has been reported that bipolar stressing has different dependencies on polarity, oxide thickness, temperature and stressing frequencies, as compared to unipolar stressing.

Liang *et al*. have reported that bipolar stressing of thin oxide leads to higher time-to-breakdown (TDDB) and charge-to-breakdown (Q_{BD}) than static stressing [4.10]. It is speculated that this was due to a combination of reduction of hole trapping and negative oxide traps generation [4.11]. Hwang *et al.* have further shown that this is only true for thicker dielectrics (> 60 Å) and the reverse is true for thinner oxides which actually show a lower TDDB for bipolar stressing than static stressing [4.12]. This result is observed to be consistent for both gate oxides and oxynitride dielectrics. Soh *et al.* further show that the unipolar stressing also exhibits significant polarity dependence with substrate-to-gate carrier injection resulting in higher TDDB than gate-to-substrate injection [4.13]. The reported data shows that TDDB for bipolar stressing in thin oxides is always higher than that of static stressing for polysilicon gate-to-substrate injection, but lower than that of substrate-to-gate injection [4.13]. While most of the reported literature have explained the various phenomena with suppression of electron trapping and enhanced hole detrapping in bipolar current stressing, Dumin *et al.* proposed that the spatial non-correlation of trap generation in bipolar stressing was the main reason for its enhanced TDDB in thick oxides (~ 80 Å) [4.14].

Using interface trap generation as a monitoring tool, Chen *et al*. have observed that interface trap generation ΔN_{IT} for unipolar stress is independent of stressing frequency, but shows frequency dependence for bipolar stress [4.15]. For bipolar stressing, frequency dependence becomes critical above 30 kHz with a linear relationship, observed between ΔN_{IT} and stressing frequency. At frequency below 30 kHz, interface trap generation is essentially independent of bipolar current stressing

frequency. Rosenbaum *et al*. have further clarified using high frequency stressing that lifetime under bipolar stress increases by a factor of 40 to 100 at frequencies above 10 kHz [4.16]. Below 10 kHz, lifetime increases linearly with frequency and is almost comparable for both bipolar and unipolar stressing.

4.2 Devices and Experimental Setup

A schematic diagram of the experimental setup as well as the principles for DCIV measurements can be found in chapter three. In this case, p+ drain is left unconnected while p+ source serves as the emitter, n+ well as the base and p-well as the collector. Interface traps were monitored via the recombination base current I_B, while oxide traps were monitored via the voltage shift in peak $I_{B,max}$. Since only p-MOSFET has this vertical p/n/p-BJT structure, only p-MOSFETs with channel area ranging from 2.5 μm^2 to 400 μm^2 were used in this study for interfacial and bulk trap measurements.

For bipolar and unipolar current stressing experimental setup, constant current pulses with alternating polarity is applied via HP4155B current source while gate voltage is monitored every 5 s. Source, drain, p-well and p-substrate are all grounded. Low pulsing frequency of 0.04 Hz is used with every change in polarity after 25 s of constant current stressing. All measurements were automated using HP IBASIC with a HP4155B semiconductor analyzer. The test devices used in this study consist of both n- and p-MOSFET fabricated on p-substrate (100) using 0.15 μm CMOS technology. Samples with channel area ranging from 1.6 μm^2 to 400 μm^2 are used while oxide thickness is 45 Å. For bipolar stressing, alternating pulses of constant current electrical stress are applied while the gate voltage is monitored. Fig. 4.1 shows a sample of the gate voltage under alternating pulses of current with current density J = 50 mA/cm^2. Low frequency pulses are selected because of the ease of implementation.

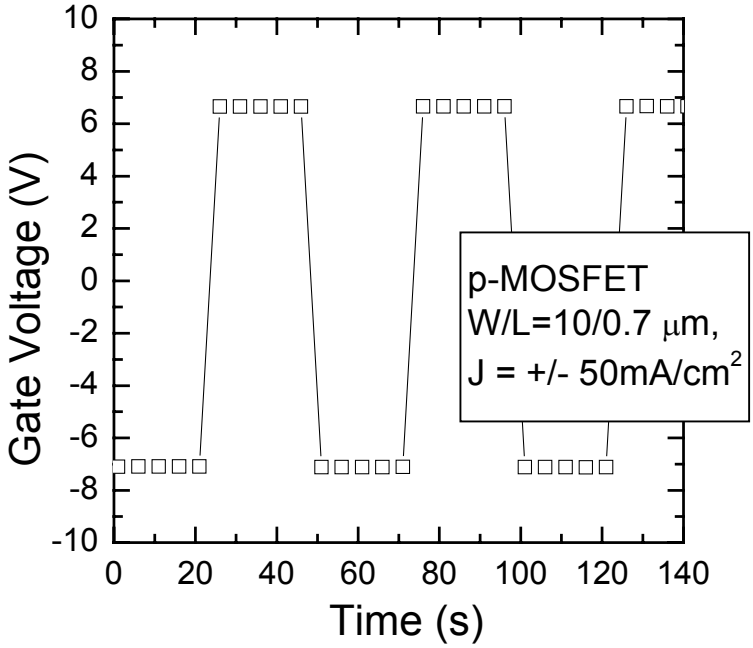


Fig. 4.1 Measured gate voltage due to application of constant bipolar current pulse of current density J = +/- 50 mA/cm^2.

4.3 Complete Evolution of Oxide Degradation Stages

In this section, bipolar current stressing was performed and its complete evolution in post-QB was observed and analyzed. A comparison of the charge-to-quasibreakdown Q_{QB} for both bipolar and unipolar current stressing is also performed and an area dependency which was not reported previously was also investigated. The full characterization of quasi-breakdown into 2 stages – recoverable and unrecoverable QB, is reported as well.

Comprehensive study of the characteristics of gate voltage and I-V characteristics at the quasi-breakdown regime suggests that there are at least 2 different stages within the quasi-breakdown. Fig. 4.2 shows the complete evolution of gate voltage under bipolar constant current stress till complete breakdown. It is observed that within quasi-breakdown, there are 2 distinct stages characterized by its electrical recoverability – recoverable and unrecoverable QB. In recoverable QB, the gate voltage is recovered to its pre-QB values at the application of the next reverse bias current pulse. This electrical recovery is only temporary and gate voltage

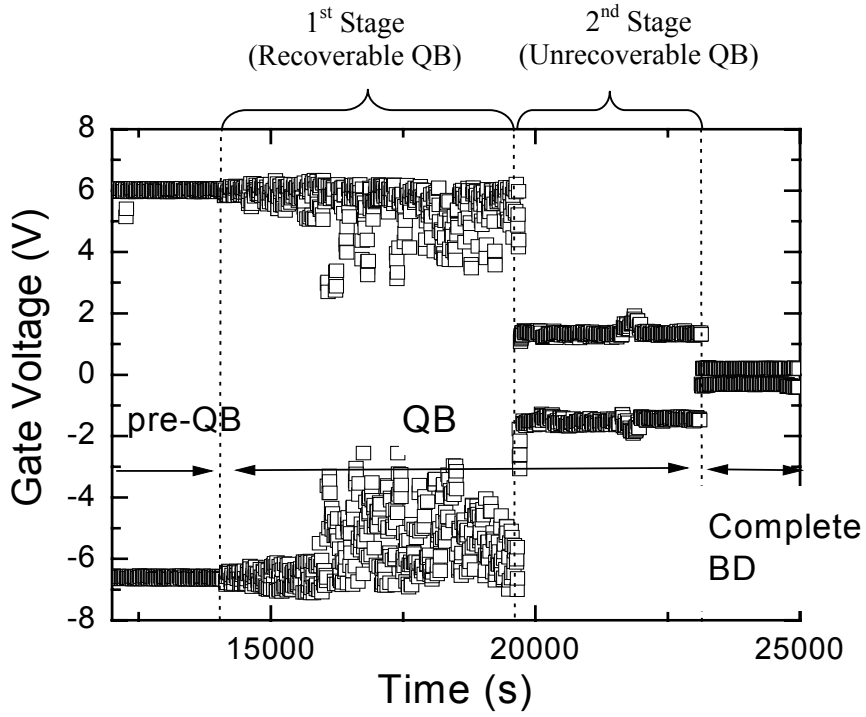


Fig. 4.2 Complete evolution of oxide degradation for thin oxide (T_{OX} = 45 Å) under bipolar constant current stress (J = +/- 10 mA/cm^2). It can be observed that within QB there are 2 stages – recoverable and unrecoverable QB. (n-MOSFET, W/L = 10/0.2 μm)

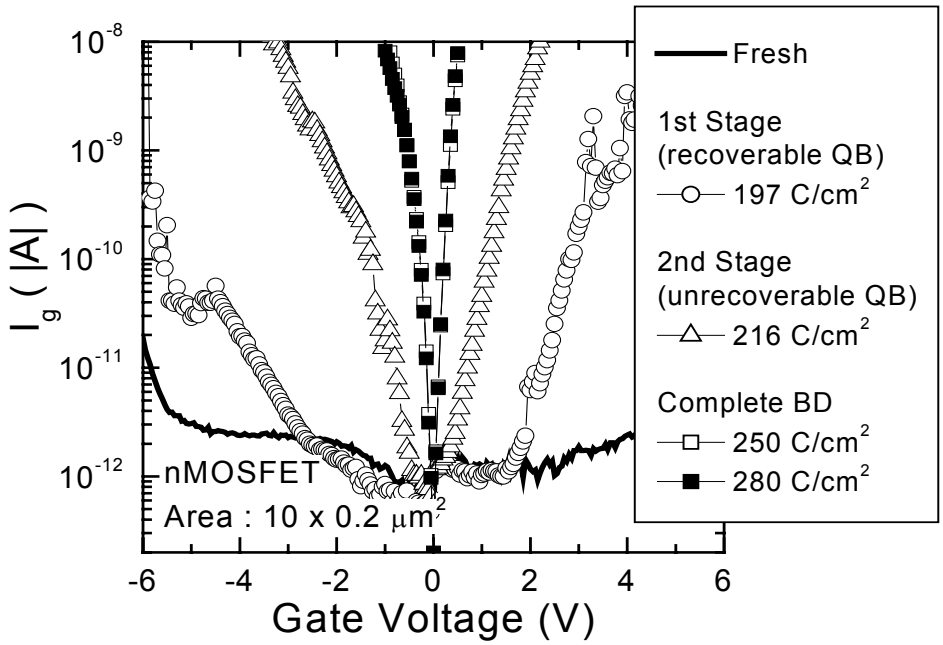


Fig. 4.3 I-V characteristics of oxides at various stages of stressing conditions – fresh, recoverable QB, unrecoverable QB and complete breakdown. (n-MOSFET, W/L = 10/0.2 μm)

constantly falls back to its post-QB values under continuous stressing. This results in a highly fluctuating gate voltage as observed in Fig. 4.2. In contrast, further stressing results in unrecoverable QB with a relatively stable gate voltage under bipolar current stressing. At this stage, no electrical recovery is observed and gate voltage is very stable, similar to complete breakdown but with much lower gate leakage current as shown by the higher voltage attained for unrecoverable QB. Fig. 4.3 shows the associated I-V curves at the various stages of degradation. At recoverable QB, I-V characteristics are constantly changing due to electrical recovery. From Fig. 4.3, it can be observed that gate leakage at recoverable QB is highly asymmetrical for positive and negative gate bias. In contrast, unrecoverable QB shows very stable and symmetrical I-V with higher leakage current at low field. Nevertheless, the I-V characteristics for unrecoverable QB is still orders of magnitude smaller than complete breakdown and can be easily distinguished as a separate stage within quasi-breakdown.

The characterization of quasi-breakdown into 2 distinct stages is further supported by the disparate response to thermal annealing. This will be covered in greater details in section 4.8 and in *Chapter 5*: *Effect of Bias and Thermal Annealing on Quasi-breakdown and its Mechanism Study*.

4.3.1 Trap generation in thin gate oxides

Since trap generation under electrical stress leads to oxide degradation, an understanding of the mechanism and its field dependencies would be critical to the total understanding of gate oxide reliability and in particular, QB. Various techniques are currently available to monitor trap generation. Among the more popular ones are charge pumping technique, direct-current current-voltage (DCIV) method and quasi-static C-V measurement. Due to its ease of use and accuracy, only the DCIV technique is used in this study.

4.4 Trap Generation and Fluence dependency

Figure 4.4 shows the peak recombination base current $I_{B,max}$ versus charge fluence for different sample areas and stressing current density. The oxide is stressed

till QB using constant current stress (CCS). Since interface traps are proportional to the recombination current I_B, it can be observed that prior to quasi-breakdown (QB), interface traps increase monotonically with the square root of charge fluence $Q^{0.52}$. At the onset of QB, $I_{B,max}$ saturates without any further increase. The value of the power exponent ~ 0.52 is similar to reported values and is consistent for the whole range of charge fluence till QB. There is also no difference for small and big channel area samples with the same power exponent of 0.52 observed for both large and small area samples.

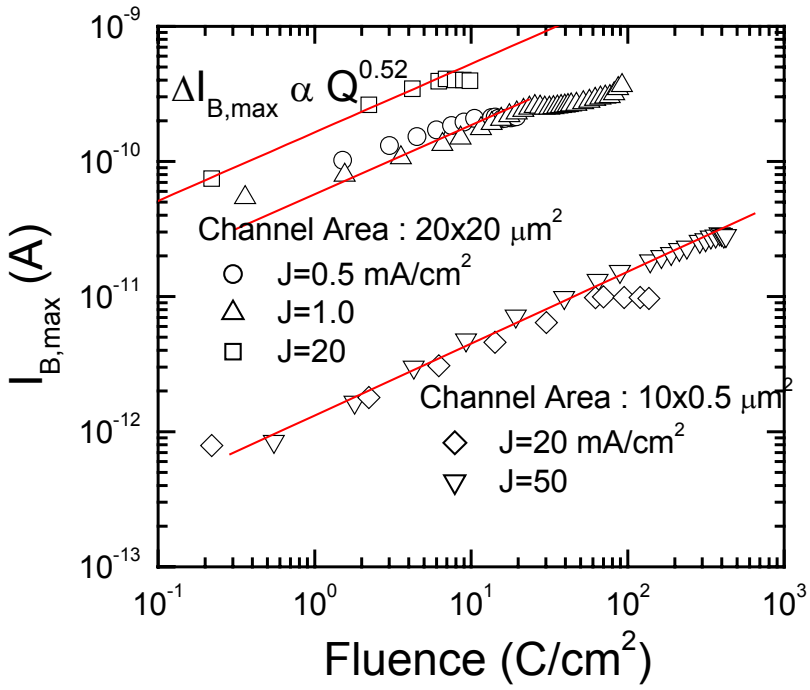


Fig. 4.4 Peak recombination current $I_{B,max}$ versus charge fluence for stressing till quasi-breakdown. Since peak $I_{B,max}$ is proportional to interface trap, it can be observed that interface trap density is proportional to charge fluence $Q^{0.52}$ independent of channel area and stressing current density.

Figure 4.5 shows the evolution of voltage shift of $I_{B,max}$, V_{GB}, which is proportional to the oxide trap density till quasi-breakdown under different CCS. Unlike interface trap generation, it can be observed that the oxide traps density has two different fluence dependencies. In the early stages, oxide trap density N_{OT} is proportional to fluence $Q^{0.043}$. As stressing is continued, oxide trap increases after 100 C/cm^2 and follows another relation $N_{OT} \alpha Q^{0.31}$. The gradient discontinuity can be interpreted as corresponding to the critical fluence whereby oxide trap generation

changes mechanism. The negative voltage shift of $I_{B,max}$ shows hole trapping or acceptor-like oxide trap formations. Holes, are generally believed to be trapped at E' center which originates from an oxygen vacancy. The E' center is a form of trivalent Si defect that results from the breaking of oxygen-deficient Si-Si strained bonds situated close to the Si-SiO_2 interface. Most literatures report a $Q^{1/3}$ dependencies which is very close to the second stage of stressing. This change in generation rates for bulk traps will be further discussed in section 4.6.

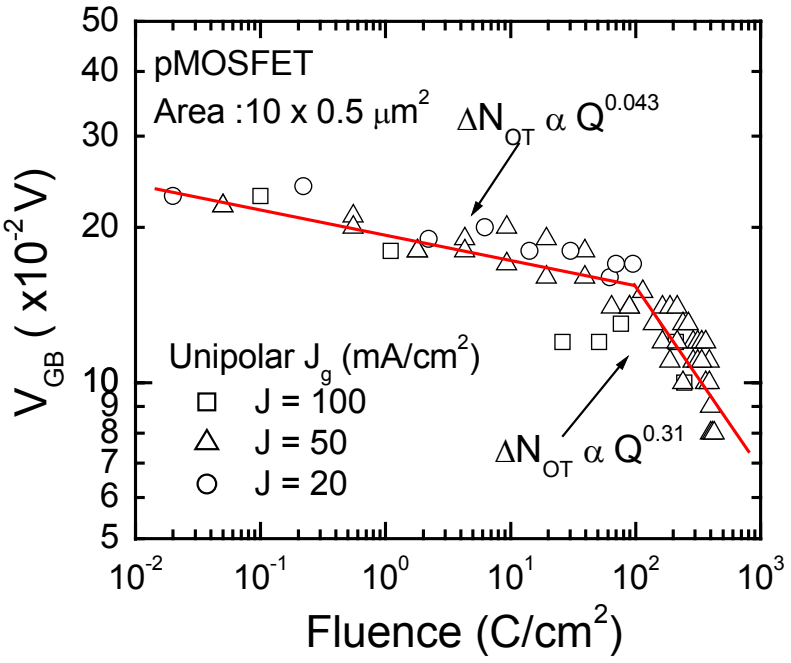


Fig. 4.5 Voltage shift of peak $I_{B,max}$, V_{GB} versus charge fluence. Since oxide trap density is proportional to V_{GB}, it can be observed that oxide trap density can be divided into 2 region. At charge fluence < 100 C/cm^2, oxide trap is proportional to $Q^{0.043}$ while beyond, oxide trap is proportional to $Q^{0.31}$.

The results on charge trapping under constant voltage stress, within the oxide bulk and at the interface prior to QB, at onset of QB and post-QB is summarized in Fig. 4.6. In this case, constant voltage stress is used for electrical stressing. It can be observed that prior to QB, both interfacial and bulk traps are increasing monotonically with stress fluency. At onset and post-QB, however, interface traps increases marginally, remaining almost constant while bulk traps, as reflected by $\Delta V_{GB,}$ shows a significant reduction till complete BD.

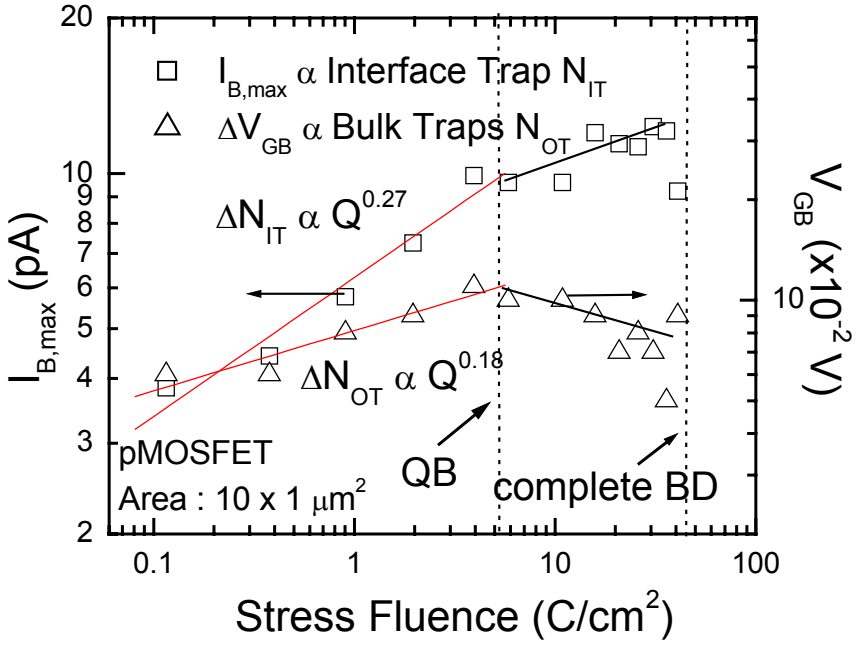


Fig. 4.6 Trap generation under constant voltage stressing. Oxide trap N_{OT} is proportional to $Q^{0.18}$ while interface trap N_{IT} is proportional to $Q^{0.27}$. (p-MOSFET, W/L=10/1 μm, V_{stress} = 6.8 V)

4.5 Critical level of trap density at onset of QB

Since trap generation has been linked to quasi-breakdown, a critical level of traps is expected at the onset of QB. Guan *et al*. have reported a constant critical level of interface traps at onset of QB showing good correlation between interface traps and oxide quasi-breakdown [4.9]. Fig. 4.7 shows the level of voltage shift (proportional to oxide trap density) at onset of QB for various electric field conditions. As the channel area of samples increases, oxide trap density is also higher, consistent with the area dependencies of the oxide trap generation mechanism. Both large and small samples show the same field dependency for oxide trap density and an increase in gate voltage (electric field) also results in higher oxide traps. This is unexpected as it shows no constant critical level of traps at onset of quasi-breakdown. Although statistical variation is expected, a general trend can be observed whereby there is an increasing oxide trap density as electric field is increased showing that the positive field dependency to critical oxide trap at QB is not coincidental.

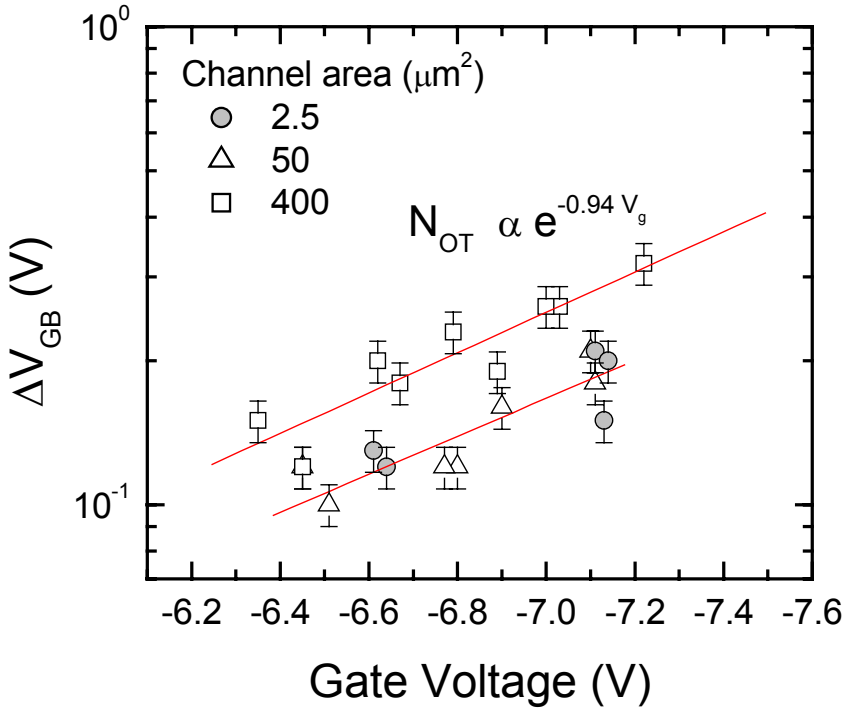


Fig. 4.7 Critical level of oxide trap at onset of quasi-breakdown for different channel area as shown. Oxides are stressed till QB using CVS at different gate bias. It can be observed that bulk traps as reflected by ΔV_{GB} at onset of QB increase with gate bias, without a single constant level expected for critical bulk defects for percolation model.

Using stress-induced leakage current (SILC) as a monitor of bulk traps, Guan *et al*. have similarly shown that SILC at onset of QB is not constant but increases with gate current density [4.9]. The results are similar to what is observed here, with varying trapped charges observed at onset of QB. The results imply that bulk traps does not reach a constant critical level at onset of QB.

Figure 4.8 shows the level of interface traps at onset of quasi-breakdown. Consistent with the data reported in [4.9], a constant level of interface traps was observed at quasi-breakdown for small and large area samples with channel area from 2.5 to 400 μm^2. This is comparable to the sample area used by Guan *et al.* which is 25 μm^2. When sample area used is small, a larger variation in the recombination current at onset of QB is observed, due mainly to the low level of recombination current at QB. Nevertheless, a constant level of recombination current was observed at onset of QB for different negative gate bias from -6.4 to -7.2 V (corresponds to more than one order of magnitude difference in gate current density).

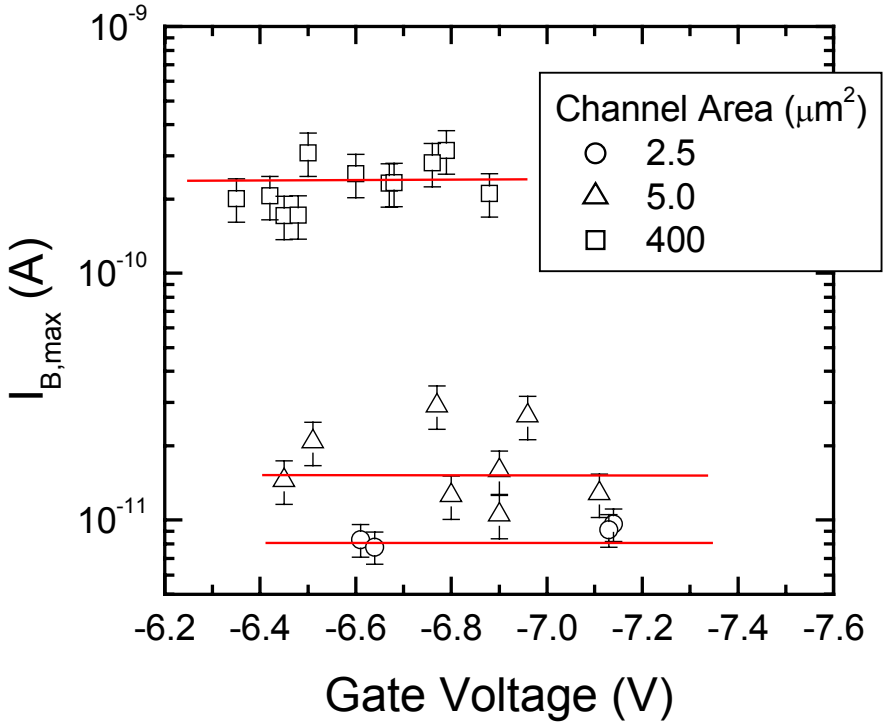
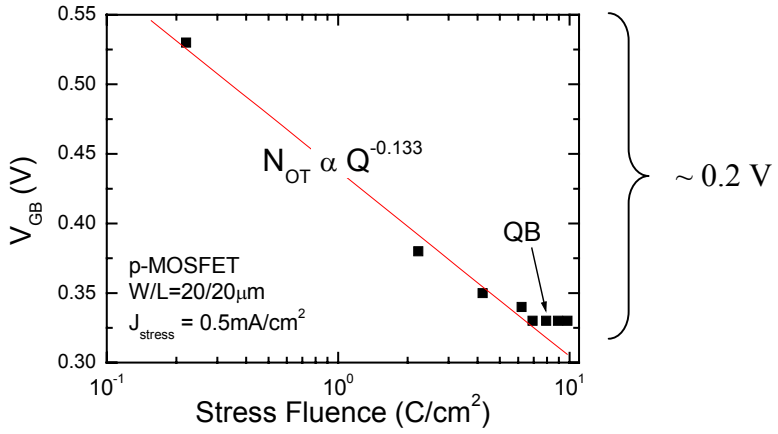


Fig. 4.8 Critical level of interface traps at onset of quasi-breakdown for different channel area. Oxides are stressed to QB using CVS at different gate bias. It can be observed that a constant level of interfacial traps is obtained irrespective of stressing gate bias. Similar results have already been reported in [4.9]

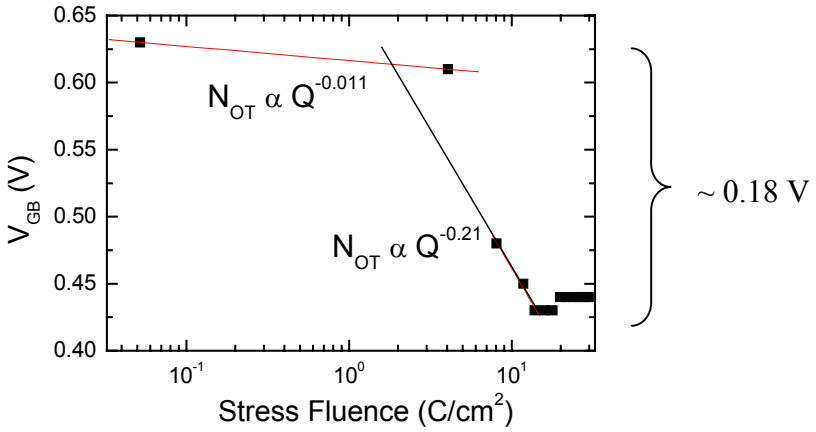
4.6 Field and Area Dependency at QB

For large area samples (~ 400 μm^2) a field dependency is observed for interface traps density at onset of QB as shown in Fig. 4.8. This lack of a constant critical level of interface traps and oxide bulk traps for large area samples at QB shows that critical interface trap density may actually have a field dependency but is masked for small area samples. From Fig. 4.5, it can be observed that oxide bulk traps generation has a kink after about 100 C/cm^2 of stressing fluence. In the initial stage of electrical stressing, oxide bulk trap generation is relatively slower at $\Delta N_{OT} \alpha Q^{0..043}$. As stressing proceeds beyond 100 C/cm^2, it is observed that oxide bulk trap generation changes to a faster rate of $\Delta N_{OT} \alpha Q^{0.31}$. The result suggests that two different types of oxide bulk traps may be formed. Fig. 4.9 shows the bulk trap generation for different stressing current density. At higher stressing current density, the first phase oxide trap generation is much more significant while the second phase

(A) $\mathbf{J_{stress} = 0.5\ mA/cm^2}$



(B) $\mathbf{J_{stress} = 2\ mA/cm^2}$



(C) $\mathbf{J_{stress} = 20\ mA/cm^2}$

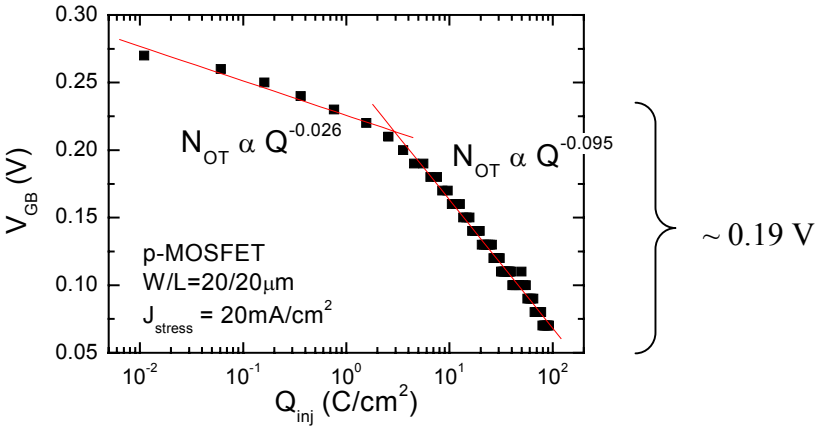


Fig. 4.9 Voltage shift of peak I_B versus stress fluence for different stressing current density (a) J_{stress} = 0.5 mA/cm^2 (b) J_{stress} = 2 mA/cm^2 (c) J_{stress} = 20 mA/cm^2. It can be observed that 2 stages exist for bulk trap generation under CCS. The crossover point is highly dependent on the magnitude of the constant current stress. In the initial stage, bulk trap generation is much slower and highly dependent on gate bias. In the second stage, trap generation is much faster and total traps generated as reflected by the magnitude of voltage shifts appear to be independent of gate bias.

of oxide trap generation remains relatively constant in terms of absolute voltage shift. As such, the field dependencies of critical bulk trap at onset of QB can be explained by the field dependencies of the initial bulk trap generation.

The second phase of oxide bulk trap generated is on the other hand, relatively constant at onset of QB. We speculate that it is this second phase of oxide traps that is responsible for QB and hence a constant critical level of oxide bulk traps will be observed at onset of QB after subtracting the effect of the initial oxide bulk traps. The results observed here, show that for both interfacial traps and oxide bulk traps, a constant critical value of traps at QB exists. In the case of interfacial traps, small and large area samples show distinctly, a constant level of interface trap at onset of QB. On the other hand, bulk traps also show a constant level, as shown by the constant voltage shift in the second stage of bulk trap generation. The results show conclusively that both a critical level of bulk and interface traps are observed at QB and the results cannot be used as conclusive evidences for interface-damage mechanism nor for linked defect path as in percolation model.

4.7 Comparison of Q_{QB} for bipolar and unipolar current stressing.

Figure 4.10 shows the charge-to-QB (Q_{QB}) for various current stressing mode using small channel area. It can be observed that in small channel area samples (< 3 μm^2), bipolar current stressing results in Q_{QB} which is orders of magnitude lower than that for negative and positive gate injection. Compared to unipolar current stressing, bipolar current stressing results in Q_{QB} which is slightly less than 1 order of magnitude lower. This difference is consistent and occurs throughout the whole range of current densities used. It was also observed that unipolar gate and substrate injection gave almost similar Q_{QB} as opposed to differing values in larger channel area. While there is significant spread in the Q_{QB} of unipolar current injection, it is observed that bipolar current injection results in very consistent Q_{QB}. The disparity in Q_{BD} results for bipolar and unipolar stressing are consistent with those obtained by Wang *et al*. using device area of 25 μm^2, who also observed lifetime decrement for

bipolar stressing compared to unipolar stress [4.19].

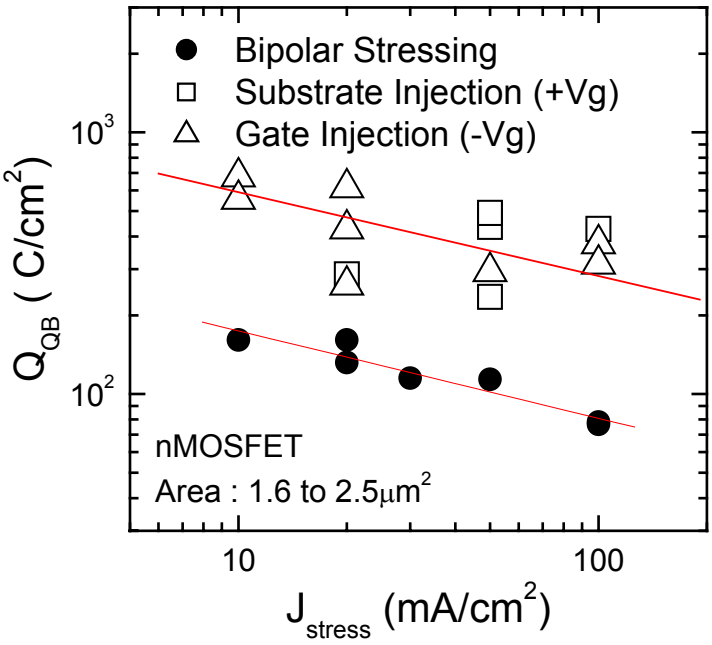


Fig. 4.10 Charge to quasi-breakdown for small channel area (< 3 μm^2) using various mode of constant current stress with different polarity injection, including unipolar and bipolar CCS. It can be observed that bipolar stressing results in much lower Q_{QB} for the same fluence as compared to unipolar stressing. Each point (differentiated by sample area) is obtained from 5-10 samples with Q_{QB} (63% values) corresponding to the zero level in the Weibull distribution.

Figure 4.11 shows the charge-to-QB, Q_{QB} for large area samples. The channel area used here is about 100 times bigger than that of the small sample area and shows the extremity in area dependency. Unlike the Q_{QB} for small channel area, large channel area samples demonstrate a Q_{QB} which is almost similar for both bipolar and unipolar gate injection current stressing. In this case, it was observed that unipolar substrate current stressing results in higher Q_{QB} which is order of magnitude higher than both bipolar and gate stressing. This is consistent with data reported in the literatures [4.13],[4.20]. In this case, gate injection appears to be the dominant breakdown mechanism which is correlated with the observation that in all the bipolar stressing breakdown at the negative current pulse first. For the same current density stress, bipolar stressing results in almost the same Q_{QB} as for negative gate bias CCS.

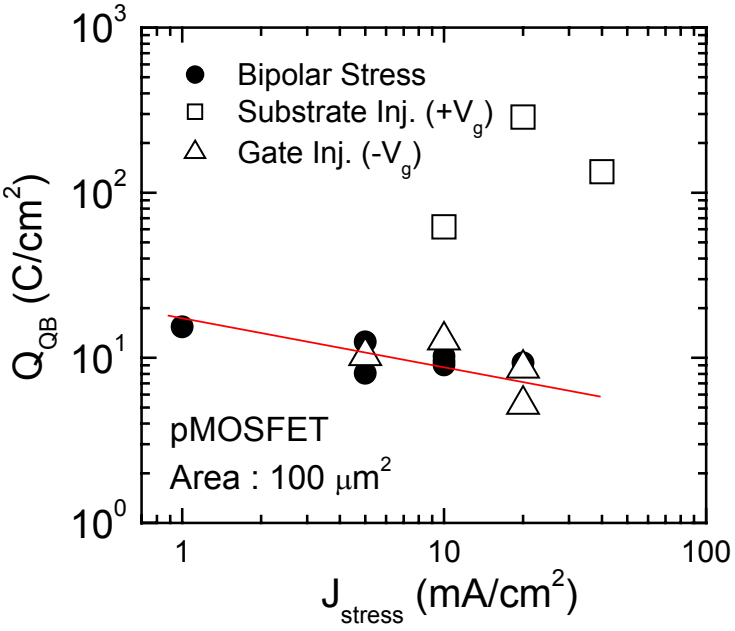


Fig. 4.11 Charge to quasi-breakdown for big channel area (~100 μm^2) using various mode of constant current injection. For large area samples, gate injection unipolar stress results in significantly lower Q_{QB} compared to substrate injection. Bipolar stressing for large samples also has low Q_{QB} and appears to be limited by the low Q_{QB} for gate injection under negative gate bias. Results are similar to [4.13]. Each point is obtained from 5-10 samples with Q_{QB} (63% values) corresponding to the zero level in the Weibull distribution.

The Q_{QB} studies show some very interesting results between small and large channel area samples. In small channel area samples, bipolar current stressing results in smaller Q_{QB} than both unipolar gate and substrate current injection while in large area samples, bipolar Q_{QB} is dominated by gate current injection. One possible explanation for lower bipolar Q_{QB} in small samples is the good spatial correlation between the traps generated by substrate and gate injection. A perfect spatial correlation will result in halving of the unipolar Q_{QB} for bipolar Q_{QB}. In the case of small area samples, we observed that bipolar Q_{QB} is less than half of that for unipolar current stressing. The enhanced degradation suggests that in addition to good spatial correlation, there are also enhanced trap generation at the potential QB spot leading to a bipolar Q_{QB} which is far lower than the combinational effect of substrate and gate injection.

In the case of big area samples, quasi-breakdown (QB) appears to be dominated by the effect of trap generation from gate injection. This is obvious comparing the different magnitude of Q_{QB} for substrate and gate injection. The lower

Q_{QB} from unipolar negative gate voltage stress (gate injection) is well documented and consistent with other data reported [4.12],[4.13],[4.21]. The result for big samples suggests that in big channel area samples, trap generation due to substrate and gate injection stressing do not have good spatial correlation, resulting in almost independent degradation from both interface with gate current stressing dominating the degradation process.

4.8 Carrier Separation Results

Figure 4.12 shows the carrier separation measurement for an unstressed p-MOSFET. For gate voltage $|V_g|> 4.5V$, well current I_w starts to increase due to F-N conduction of valence band electron injection from the gate. The onset of the rise in gate and well current corresponds to the valence band barrier height (~ 4.3 eV) and shows that gate current at high field is due to valence-band electron injection from the p+ gate. Above $|V_g| > 5V$, hole current as shown by source current I_s changes the sign, showing onset of impact ionization which results in increased electron-hole pair generation. For $|V_g| > 5$ V, gate, source and well current are related as follow:

$$I_w = I_g + I_s$$
$$I_s = \gamma I_g \tag{4.1}$$

where γ is the impact ionization or quantum yield factor and is almost unity.

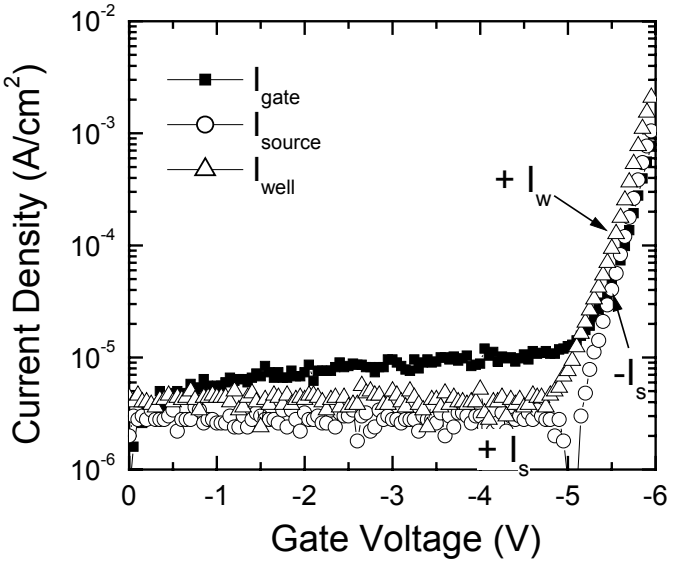


Fig. 4.12 Carrier separation measurement showing gate, source and substrate current component at fresh, unstressed state. (p-MSOFET, T_{ox} = 45 Å, W/L = 10/0.5 μm)

Figure 4.13 shows the carrier separation measurement at onset of QB under low stressing current density (J_{stress} = 10 mA/cm^2) CCS conditions. It can be observed that at onset of QB, gate current increases at low field resulting in significant leakage current. For $|V_g|$> 3 V and below 5.5 V, both well and gate current are almost equal $I_W \approx I_g$ showing that there is minimal impact ionization. At the same time, hole current is observed but this is not dominant. Above V_g > 5.5 V, hole current reverses the sign due to significant impact ionization of the injection electron and excessive holes generated are removed from the source through the external ground. Studies have reported the dominance of hole conduction [4.22] at QB but have not reported that of electron dominance in the initial stage of QB. Fig. 4.14 shows the carrier separation measurement after additional post-QB fluency of 38 C/cm^2. At this stage, hole current from substrate dominates and is far higher than the electron current even at low field. The gate leakage current $I_g = I_s$ at both low and high field showing that gate leakage current is due primarily to the hole direct tunneling (DT) from substrate. This conforms to the result of other studies [4.22] and is the commonly observed phenomenon. It was observed from bias anneal studies that the oxides are still in

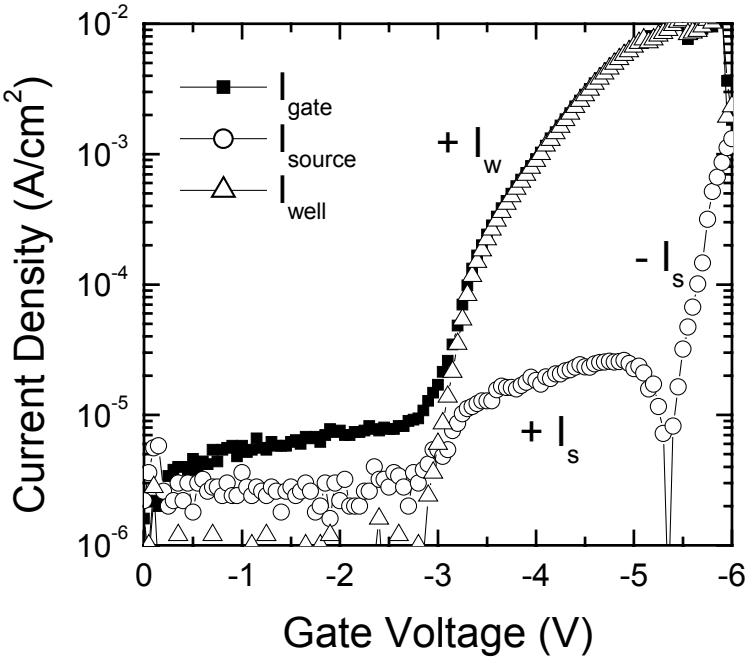


Fig. 4.13 Carrier separation measurement at onset of QB, which is attained after 222C/cm^2 of electron fluency. Sample is still in recoverable QB stage and is the same one used in Fig. 4.12.

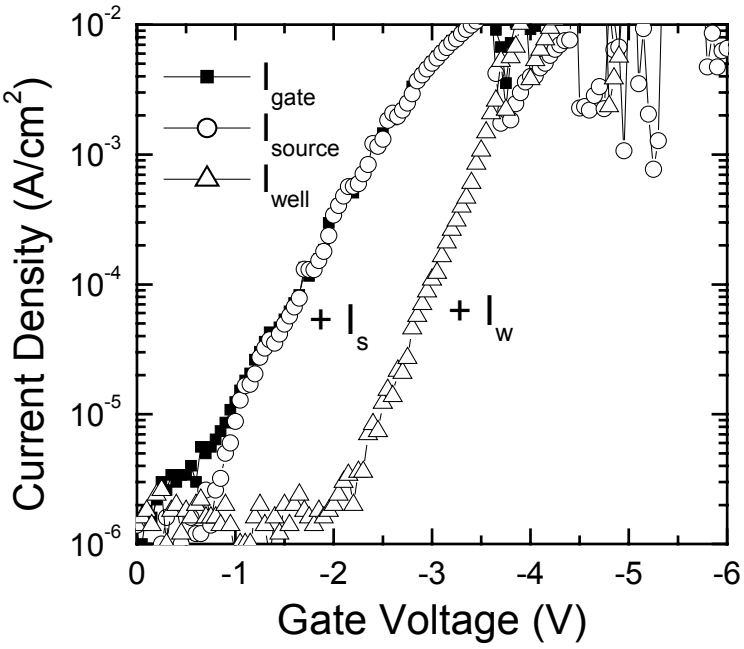


Fig. 4.14 Carrier separation measurement after post-QB stress (additional electron fluency of 38 C/cm^2) within recoverable QB stage. Sample used is the same as Fig. 4.12

recoverable QB at this stage. Using carrier separation, Uno *et al*. have also shown similar results for the initial stage recoverable QB where electrons are dominant as shown in Fig. 4.13 with slight hole component [4.23]. Calling it Pre-breakdown (PreBD), this phenomenon is clearly distinguished from SILC which only contains an electron tunneling component.

Figure 4.15 shows the carrier separation measurement of an oxide when stressed till unrecoverable QB. It can be observed that the leakage current of source, well and gate are all very similar to that of Fig. 4.14 in the recoverable QB stage. Hole and electron leakage current as shown by source I_S and well I_W current respectively are much higher even at low field as compared to that in the recoverable QB but otherwise retain its general shape. Moreover, it can be observed that hole DT current starts to rise at gate voltage $V_g = 0$ V whereas electron DT current appears only at $V_g \approx 1$ V, corresponding to the field necessary to raise the valence band in the p+ gate above the conduction band of the n+ substrate (≈ Si bandgap barrier of 1.12 eV for a p+ gate/p-MOSFET/n+ well).

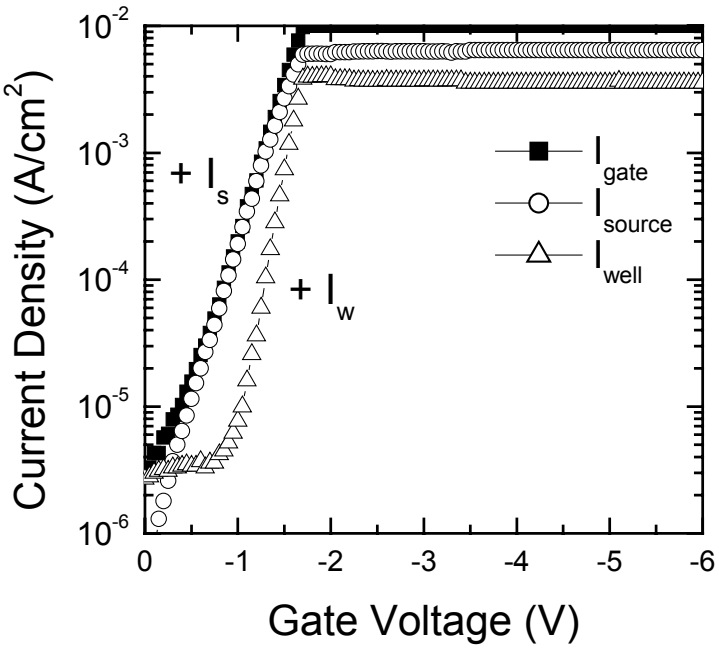


Fig. 4.15 Carrier separation measurement after post-QB stress but stressed to unrecoverable QB (p-MOSFET, T_{ox} = 45 Å, W/L = 10/1 μm, J_{stress} = 10 mA/cm^2)

To determine the mode of carrier transport across the gate oxide, the impact ionization quantum yield γ can be used to determine the energy of the injected electron at the SiO_2/Si interface [4.7],[4.24]. However since the method measures the impact ionization quantum yield, it is applicable only in the high electric field, which narrows its applications to high field analysis. Fig. 4.16 shows the quantum yield γ versus gate voltage for fresh, QB and post-QB states. At the unstressed state, it can be observed that significant impact ionization occurs for $|V_g| > 5$ V with unity quantum yield $\gamma \approx 1$. At onset of QB, quantum yield γ drops at low field due primarily to increase in electron current without corresponding increase in impact ionization.

Above $|V_g| > 5$ V, quantum yield γ starts to increase but is still significantly lower than that of unstressed state. Since quantum yield of impact ionization γ is strongly dependent on electron energy, it is possible to calculate any energy loss of the electron during the transport across the gate oxide by correlating the quantum yield at fresh and QB state. From Fig. 4.16, it can be observed that the quantum yield at QB is much lower than that of unstressed state at the same gate voltage. The results suggest that there is significant energy loss for electron during transport across the gate oxide at onset of QB.

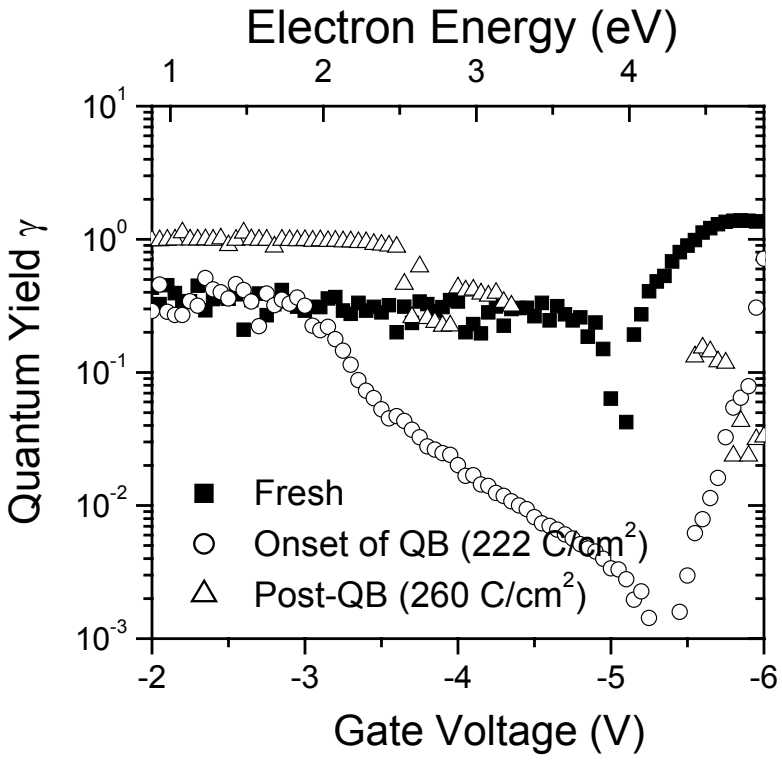


Fig. 4.16 Experimental quantum yield as a function of gate voltage and resulting electron energy for unstressed gate oxides and post-QB oxides as shown inset.

4.9 F-N and Direct Tunneling Modeling Equations

Using carrier separation, it was observed that the hole and electron current-voltage (I-V) curves have distinctively different shapes at QB and post-QB. While most studies have attempted to model the gate leakage current, it may be more appropriate to model the substrate and well current since gate leakage current is a sum of these two components at low field.

4.9.1 Electron Leakage Current

In the case of p-MOSFET under inversion conditions, electron conduction monitored through well current was observed to be parallel to the fresh Fowler-Nordheim (F-N) I-V curves. Fig. 4.17 shows the electron current (well current I_{well}) for p-MOSFET under inversion at fresh and at onset of QB. Due to charge trapping within the gate dielectrics, the electric field within the gate dielectric is modified. At QB, a critical level of charge trapping within the dielectric anode is reached and this is emulated by the term V_{shift}. In this case, V_{shift} represents the modification of electric

field within the silicon dioxides, due to hole trapping at the anode at the onset of QB. Using various parallel shift, a best fit was observed for V_{shift} = 2.2 V when applied to fresh F-N current and fitted to the well current at QB. The term LDR (localized damage region) model was invoked to represent charge trapping resulting in quasi-oxide damage at the anode. This results in a modification of the electric field within the gate dielectrics and is similar to adding an additional V_{shift} term to the dielectric internal electric field. Varying the energy barrier ϕ_B and oxide thickness T_{ox} did not give a well-fitting curve to the electron current at QB. The effective F-N tunneling current can then be modeled using the classical F-N tunneling equation as shown in (4.2) but with a modified V'_{OX} where $V'_{Ox} = V_{OX} + V_{shift}$ and $E'_{ox} = V'_{ox} / T_{OX}$

$$
\begin{aligned}
J &= CE_{ox}^{2} \exp\left(-\frac{\beta}{E'_{ox}}\right) \\
C &= \frac{q^3}{16\pi\hbar}\left(\frac{m^*}{m}\right)\left(\frac{1}{\Phi_b}\right) = 1.54x10^{-}6\left(\frac{m^*}{m}\right)\left(\frac{1}{\Phi_b}\right)\left[A/V^2\right] \\
\beta &= -\frac{4(2m^*)^{1/2}\Phi^{3/2}}{3\hbar q} = 6.83x10^{7}\left(\frac{m^*}{m}\right)^{1/2}\Phi^{3/2}\left[V/cm\right]
\end{aligned}
\tag{4.2}
$$

Although most of the electron current at QB can be fitted using this method with the appropriate V_{shift}, it was also observed that certain electron current have I-V characteristics similar to that as shown in Fig. 4.18 at QB which is not amenable to fitting using the F-N tunneling equation. Using further stressing, it was observed that electron current can be recovered to I-V characteristics that can be easily fitted with the F-N tunneling current equation (4.2) with an appropriate V_{shift}. This recovery in well current due to continual stressing is slightly dissimilar to the recovery observed using reverse bias anneal. One significant difference is the magnitude of reduction in leakage current. In the case of reverse bias annealing, almost complete recovery to SILC is observed using low reverse bias fluence (~ 2.4 C/cm^2) while continual stressing did not result in significant reduction in leakage current even after high stress fluence (~ 238 C/cm^2). The recovery to F-N like conduction after post-QB continual stressing can be explained by electron compensation of deep level trapped holes resulting in a graded level of trap energy which revert the oxide back to F-N dependence direct tunneling of electrons.

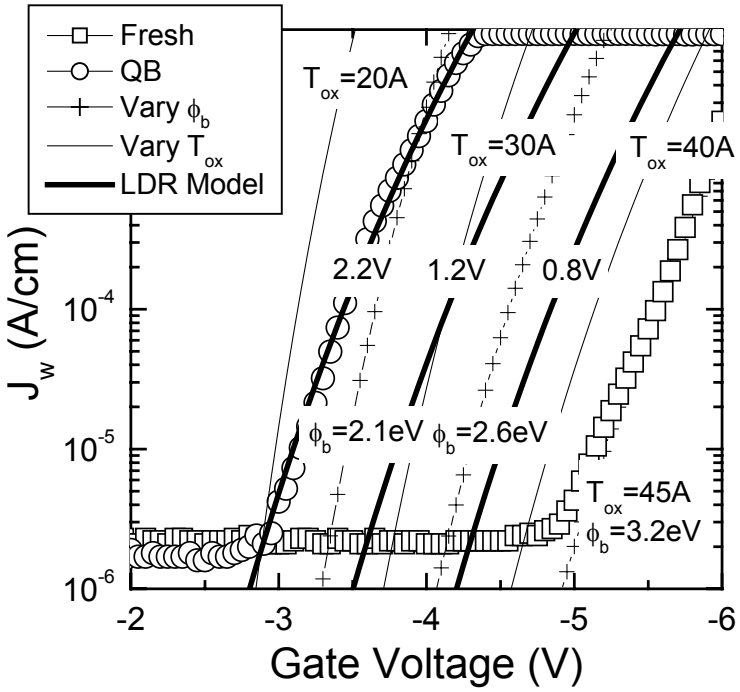


Fig. 4.17 Carrier separation for well current component at fresh and onset of QB state. F-N current is simulated using (3.5) with varying energy barrier, oxide thickness and electric field shift. Good fit observed for experimental I_{well} and F-N current using electric field shift V_{shift} = 2.2 V.

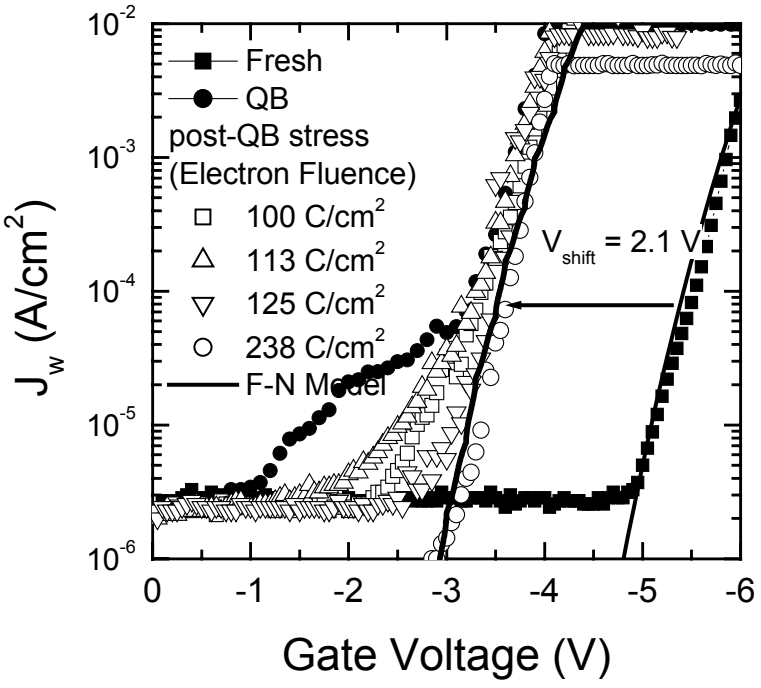


Fig. 4.18 Evolution of well current component for post-QB stage under continual stressing. Good fit observed for well current component with simulated F-N tunneling current. (T_{ox} = 45 Å, W/L = 10/0.7 μm, p-MOSFET)

4.9.2 Hole dominance leakage current after QB

At subsequent stages of QB, hole dominance over electron become the main leakage current component. At this stage, it was observed that hole current at post-QB has an I-V characteristic, which can be described by the direct tunneling current equation. Fig. 4.19 shows the hole current (source current) for p-MOSFET at post-QB where source current has become dominant. Using the direct tunneling equation (4.2), it was observed that hole current as shown by source current component can be fitted relatively well with the DT modeling current equation as shown in Fig. 4.19 using an effective oxide thickness of T_{ox} = 27 Å.

$$J_{DT} = J_{FN} \times 2\left(1 - \frac{V_{Ox}}{2\Phi_B}\right)\left(\frac{\Phi_B}{V_{OX}}\right)\exp\left(\frac{\beta\left(\frac{\Phi_B - V_{OX}}{\Phi_B}\right)^{3/2}}{E_{OX}}\right) \tag{4.3}$$

The result suggests that in the early stage of recoverable QB where electron current is dominant, direct tunneling with F-N dependencies occurs while in the subsequent stage, direct tunneling (DT) of holes becomes dominant. Based on the

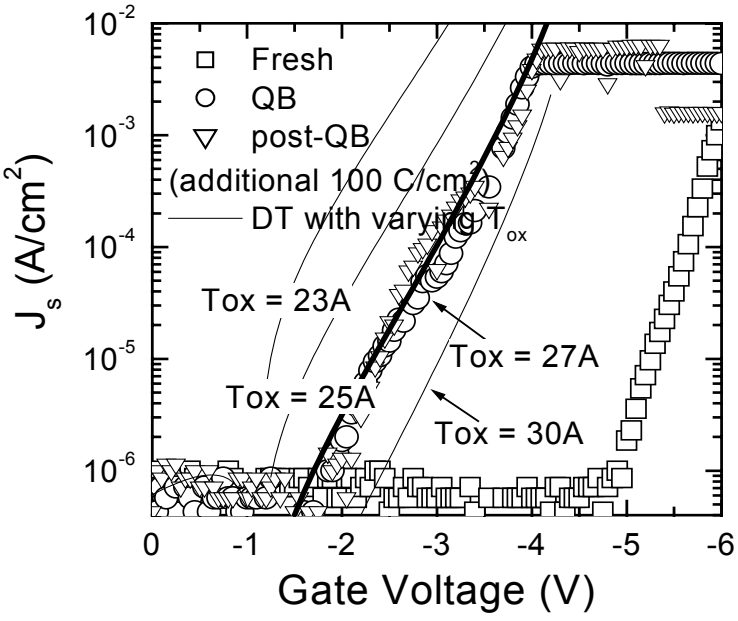


Fig. 4.19 Evolution of source current component after QB in hole dominant regime. Relatively good fit observed between experimental data at QB and direct tunneling current modeling using (3.6). (T_{ox} = 45 Å, W/L = 10/0.7μm, p-MOSFET)

simulation fittings, the effective oxide has been reduced by about 18 Å from the initial oxide thickness of 45 Å. This is reasonable and is almost similar to the oxide thinning parameter used in [4.1]. The results also explain the initial controversy regarding the DT model due to the 'unrealistic' barrier height used in simulation. Based on best fit, a barrier height of 4.2 eV was found [4.25] for post-QB fitting with DT model which was rejected as unrealistic. However, it is clear that holes instead of electrons governed the subsequent stages of QB and modeling should be performed with holes DT current instead of electrons. The results also explain why quasi-breakdown is observed predominantly in thin oxides where it is easier to approach the direct tunneling thickness regime.

4.10 Proposed Model for QB Mechanism

The bipolar stressing experiment in section 4.3 has shown that there exists two stages within QB – recoverable and unrecoverable QB. From the thermal and bias annealing studies performed [4.26], it has been shown that strong evidences point to the possibility that oxide bulk traps lead to QB leakage current. These results will be further elaborated in the next chapter. Voltage shift in the DCIV measurement has shown that these are positively charged, with strong possibility of holes trapping within the oxide. From the carrier separation measurement, it was further shown that electron dominance exists in the initial phase followed by hole dominance within recoverable QB. Impact ionization quantum yield studies has shown that electron energy relaxation occurs during transport across the gate oxide after QB, similar to SILC [4.7]. With these cumulative evidences, a simple model of hole trapping at the anode is proposed.

Figure 4.20 shows holes trapping at the anode due to Fowler-Nordheim (F-N) electron injection from cathode (gate). F-N electron injection causes impact ionization at the SiO_2/Si interface under high electric field resulting in the formation of both electrons and holes which are collected at the substrate and source/drain respectively. Some holes are injected into the gate oxides and are trapped predominantly near the SiO_2/Si interface due to the low mobility of holes. Reverse bias annealing experiment

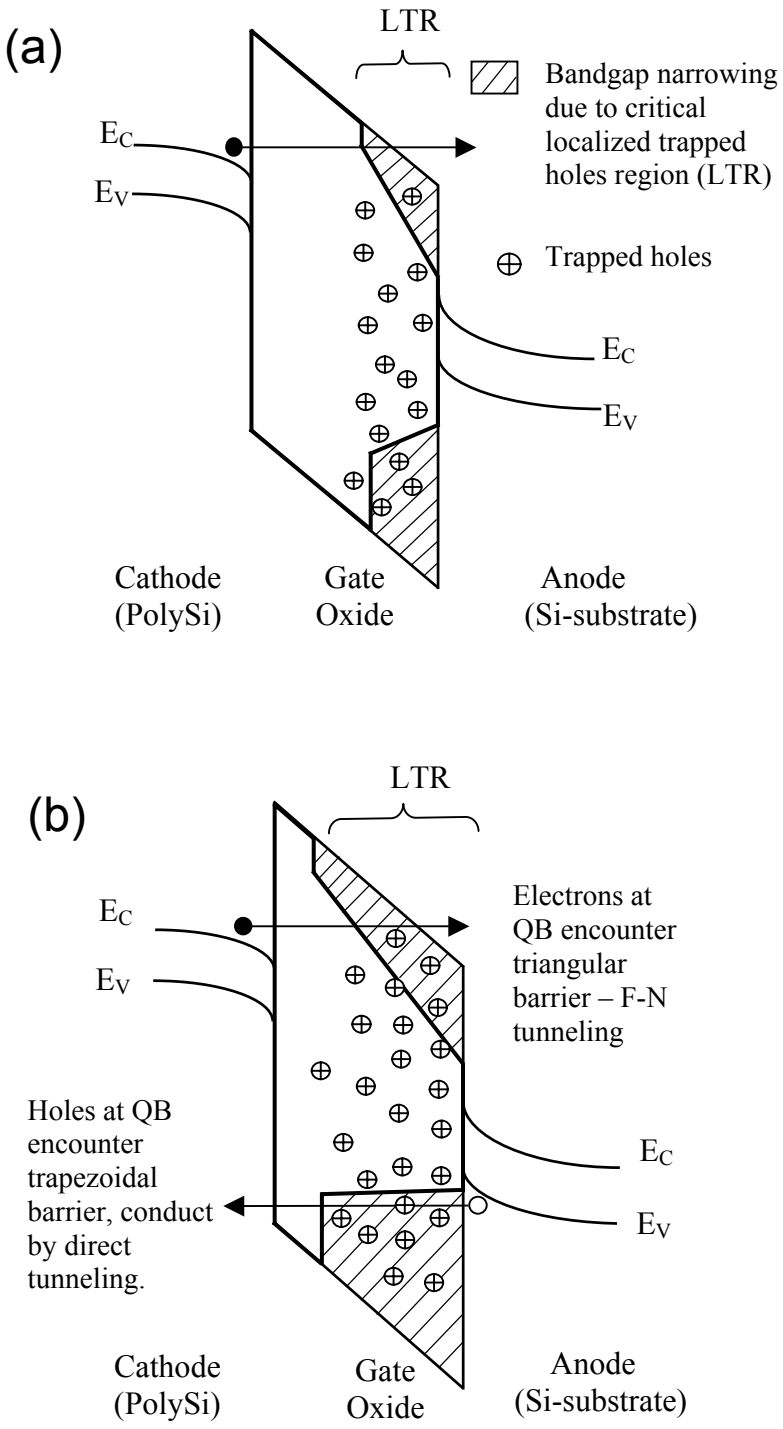
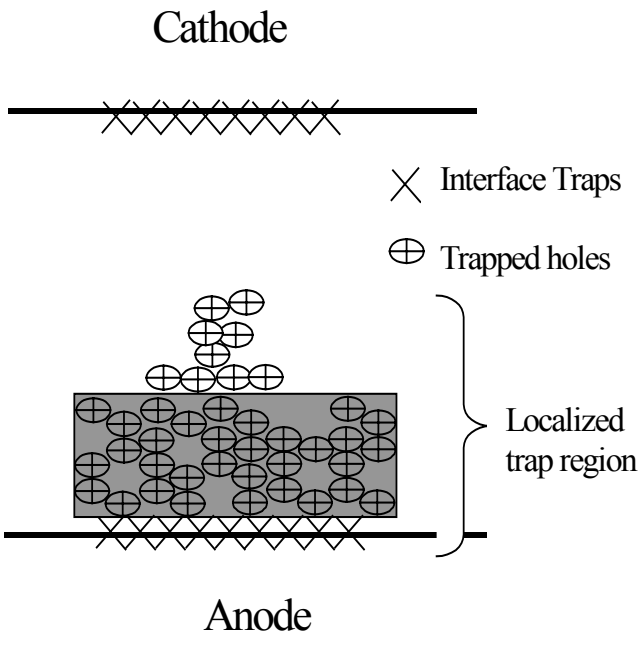


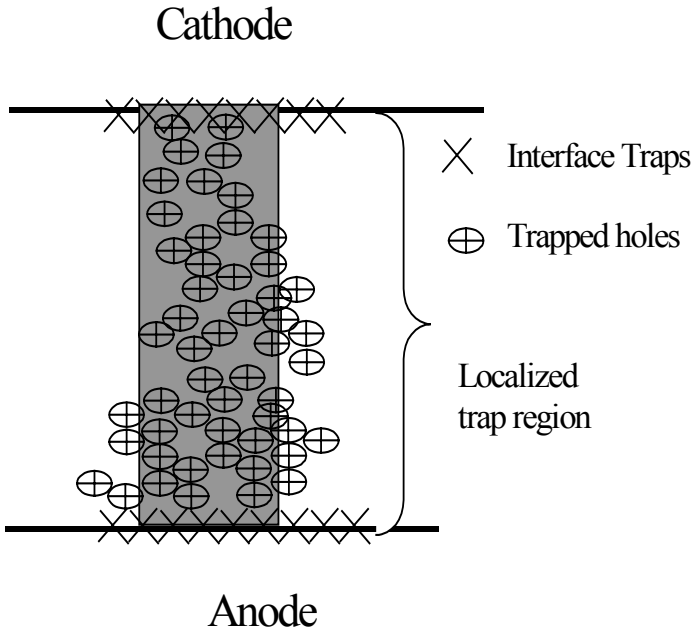
Fig. 4.20 A schematic drawing of energy band diagram for localized trap region (LTR) model. Hole trapping at anode results in distortion of energy band and formation of localized trap region (LTR) causing bandgap narrowing near the anode. (a) In initial stage, electron conduction by F-N tunneling. (b) Further stressing extends the LTR, resulting in hole direct tunneling. Electrons and holes are indicted by solid and open circles, respectively. .

suggests that the spatial distribution of the traps is not uniformly distributed across the whole oxide but resides predominantly near the SiO_2/Si interface as opposed to that of percolation model [4.2]. Since stress-induced leakage current (SILC) has also been attributed to trapped holes [4.27],[4.28], QB can be considered as a continuation of the SILC trapping mechanism with some significant differences : Trapped holes in QB exist in sufficient quantities to form a localized trap region, reducing the effective oxide thickness and hence allowing carriers to tunnel through by direct tunneling (DT). Fig. 4.20 shows the hole and electron transport at onset of QB. Essentially, both electrons and holes are transported across the oxide by direct tunneling at onset of QB. However due to the graded distribution of trapped holes energy level, it can be expected that electron transport will have a modified Fowler-Nordheim (F-N) conduction while holes will encounter a primarily direct tunneling conduction.

Figure 4.21 shows the conduction mechanism at recoverable and unrecoverable QB. At onset of QB, which is usually characterized by electrical recovery, the localized trap region (LTR) formed by deep hole traps, are mainly localized near the anode. At this stage, either electron or hole dominance is possible with carrier conduction primarily by direct tunneling and a modified F-N conduction mechanism for holes and electrons respectively. Further stressing extends the LTR into the oxide eventually linking both the anode and cathode with trapped holes. Once this occurs, unrecoverable QB is attained and conduction proceeds by a direct conduction path similar to the percolation path model.



(A) Recoverable Quasi-breakdown (QB)



(B) Unrecoverable Quasi-breakdown (QB)

Fig. 4.21 Schematic illustration of evolution of the localized trap region (LTR) formed by deep level trapped holes at various stages of QB (A) at onset of QB, within recoverable QB, LTR is mainly localized at the anode and conduction proceed by direct tunneling of both holes and electrons. (B) at unrecoverable QB, LTR has extended the whole oxide forming a direct conduction path.

4.11 Summary

In this chapter, the effect of bipolar current stressing on quasi-breakdown is studied. An area dependency of charge-to-quasibreakdown Q_{QB} is observed for bipolar stressing and this is attributed to spatial correlation of the trap generation due to the different polarity stressing. Using bipolar current stressing, it was further observed that quasi-breakdown can be characterized into 2 distinct stages – recoverable and unrecoverable QB. In recoverable QB, gate leakage current can be recovered to pre-QB level by the application of a reverse bias anneal. In unrecoverable QB, no electrical recovery is observed and current-voltage (I-V) characteristics are much more stable with higher gate leakage current.

Using the DCIV measurement technique, trap generation under F-N constant current stressing is observed. It was observed that oxide bulk traps have two distinct types – field dependent and field independent traps. Trap generation data was also found to correlate with reported literature for both interface and the field independent oxide bulk trap generation.

The conduction mechanism at onset of quasi-breakdown (QB) and after subsequent post-QB stress was also investigated. It was observed that the post-QB leakage current evolves from an electron dominance to hole dominance, all within recoverable QB. The electron leakage current at QB can be described by a modified Fowler-Nordheim (F-N) relation with electric field lowering due to positive hole trapping. On the other hand, the hole leakage current can be very well described with a direct tunneling relation. A simple model of hole trapping at the anode is proposed and QB occurs due to the formation of a localized trap region (LTR) resulting in effective thinning of the gate oxide. Recoverable QB is explained by the localization of LTR to the region near the anode, resulting in direct tunneling of both electrons and holes. Continual stressing extend the LTR eventually linking the anode and cathode forming a direct conduction path and triggering unrecoverable QB. The model is able to explain the various phenomena observed so far and is consistent with observations of the direct tunneling model first proposed by Lee and Cho *et al* [4.1].

Reference

[4.1] S. H. Lee, B. J. Cho, J.C. Kim, and S.H. Choi, "Quasi-breakdown of ultra thin gate oxide under high field stress," *IEDM Tech. Dig*., pp. 605-608, 1994.

[4.2] M. Depas. T. Nigam, and M. M. Heyns, "Soft breakdown of ultra-thin gate oxide layers," *IEEE Trans. Electron Dev*., vol. 43, pp.1499-503, 1996.

[4.3] K. Okada and K. Taniguchi, "Electrical Stress-induced Variable Range Hopping Conduction in Ultra thin Silicon Dioxides," *Appl. Phys. Lett*., vol. 70, pp. 351 – 353, 1997.

[4.4] J. Sune and E. Miranda, "Post Soft Breakdown Conduction in SiO_2 Gate Oxides," *IEDM Tech. Dig.,* pp. 533-536, 2000.

[4.5] H. Guan, B. J. Cho, M. F. Li, Y. H. He, Z Xu, and Z. Dong, "A study of Quasi-breakdown Mechanism in Ultra-thin Gate Oxide by using DCIV Technique," *7th Proc. IPFA*, pp. 81- 84, 1999.

[4.6] A. Halimaoui, O. Briere, and G. Ghibaudo, "Quasi-breakdown in Ultrathin Gate Dielectrics," *Microelectronics Engrg*., vol. 36, pp. 157-160, 1997.

[4.7] J. Sune and E. Miranda, "Are Soft Breakdown and Hard Breakdown of Ultrathin Gate Oxdies Actually Different Failure Mechanism?," *IEEE Electron Dev. Lett.,* vol. 21, no. 4, pp. 167-169, 2000.

[4.8] T. Tomita, H. Utsunomiya, T. Sakura, Y. Kanakura, and K. Taniguchi, "A new soft breakdown model for thin thermal SiO_2 film under constant current stress," *IEEE Trans.Electron Dev*, vol. 46, pp. 159-165, 1999.

[4.9] H. Guan, B. J. Cho, M. F. Li, Z. Xu, Y. D. He, and Z. Dong, "Experimental Evidence of Interface-Controlled Mechanism of Quasi-Breakdown in Ultra-thin Gate Oxide," *IEEE Electron Dev. Trans.,* vol. 48, no. 5, pp.1010-1013, 2001.

[4.10] M. S. Liang, S. Hadda, W. Cox, and S. Cagina, "Degradation of very thin gate oxide MOS devices under dynamic high field/current stress," *IEDM Tech. Dig.*, p. 394, 1984.

[4.11] Y. Fong, I. C. Chen, S. Holland, J. Lee, and C. Hu, "Dynamic stressing of thin oxides," *IEDM Tech, Dig.,* p. 664, 1986.

[4.12] H. Hwang and J. Lee, "Anomalous Breakdown Behaviour in Ultrathin Oxides and Oxynitrides under Dynamic Electrical Stress," *IEEE Electron Dev. Lett*., vol 13, no. 9, pp. 485-487, 1992.

[4.13] S. H. Soh, M. Naidu, J. Hwu, and L. Sadwick, "Thickness and Temperature Effects on TDDB for DC and Dynamic Stressing of Thin Oxide," *Int. Reliab. Workshop*, pp. 153-158, 1996.

[4.14] D. J. Dumin and S. Vanchinathan, "Bipolar Stressing, Breakdown, and Trap Generation in Thin Silicon Oxides," *IEEE Trans. On Electron Dev*., vol. 41, no. 6, p.936-939, 1994.

[4.15] T. P. Chen, S. Li, S. Fung, and K. F. Lo, "Interface trap generation by FN Injection under Dynamic Oxide Field Stress," *IEEE Tran. Electron Dev*., vol. 45, no. 9, pp. 1920 – 1926, 1998.

[4.16] E. Rosenbaum and C. M. Hu, "High-Frequency Time-Dependent Breakdown of SiO_2," *IEEE Electron Dev. Lett.,* vol. 12, no. 6, pp. 267-269, 1991.

[4.17] A. Neugroschel, C. T. Sah, M. Han, M. S. Carroll, T. Nishida, J. T. Kavalieros, and Y. Lu, " Direct-current Measurements of Oxide and Interface Traps on Oxidized Silicon," *IEEE Trans on Electron Dev*., vol. 42, pp. 1657-1662, 1995.

[4.18] B. B. Jie, M. F. Li, C. L. Lou, W. K. Chim, D. S. H. Chan, and K. F. Lo, " Investigation of interface traps in LDD p-MOSFETs by DCIV method," *IEEE Electron Dev. Lett*., vol. 18, pp. 583-585, 1997.

[4.19] B. Wang, J. S. Suehle, E. M. Vogel, and J. B. Bernstein, "The effect of stress interruption and pulsed biased stress on ultra-thin gate dielectric reliability," *Int. Reliab. Workshop*, pp. 74-79, 2000.

[4.20] G. Q. Lo, S. Ito, D. L. Kwong, V. K. Mathews, P. C. Fazan, and A. Ditali, "Charge Trapping/Detrapping and dielectric breakdown in $SiO_2/Si_3N_4/SiO_2$ stacked layers on rugged poly-Si under dynamic stress", *30th Annual Int. Reliability Physics Symp.,* pp. 42-48, 1992.

[4.21] E. Rosenbaum, Z. Liu, and C. Hu, "The effects of oxide stress waveform on MOSFET performance," *IEDM Tech. Dig.,* pp. 719-722, 1991.

[4.22] H. Guan, M. F. Li, Y. He, B. J. Cho, and Z. Dong, "A thorough study of quasi-breakdown phenomenon of thin gate oxides in dual-gate CMOSFET's," *IEEE Trans. Electron Dev*., vol. 47, no. 8, pp. 1608-1616, 2000.

[4.23] S. Uno, A. Ishida, K. Deguchi, Y. Kamakura, and K. Taniguchi, "Carrier separation measurement of leakage current under prebreakdown in ultrathin SiO_2 films," *J. Appl. Phys*., vol. 89, pp. 8336-8338, 2001.

[4.24] A. Ghetti, M. Alam, J. Bude, D. Monroe, E. Sangiorgi, and H. Vaidya, "Analysis of trap-assisted conduction mechanisms through silicon dioxide films using quantum yield," *IEDM Tech. Dig.*, pp. 723-726, 1999.

[4.25] M. Houssa, T. Nigam, P. W. Mertens, and M. M. Heyns, " Model for the current-voltage characteristics of ultra thin gate oxides after soft breakdown," *J. Appl. Phys.,* vol. 84, pp. 4351-4355,1998.

[4.26] W. Y. Loh, B. J. Cho, and M. F. Li, " Investigation of Quasi-breakdown Mechanism through Post-QB Thermal Annealing," *Jpn. J. Appl. Phys.*, vol. 41, pp. 2873-2877, 2002.

[4.27] J. Maserjian and N. Zamani, "Observation of positively charged state generation near the Si/SiO_2 interface using Fowler-Nordheim tunneling," *J. Vac. Sci Tech.,* vol. 20, pp. 743-748, 1982.

[4.28] N. K. Zous, T. Wang, C. C. Yeh, C. W. Tsai, and C. Huang, "Transient effects of positive oxide charge on stress-induced leakage current in tunnel oxides," *Appl. Phys. Lett.,* pp. 743-736, 1999.

Chapter 5

Effect of Bias and Thermal Annealing on Quasi-breakdown and its Mechanism Study

5.1 Introduction

In the previous chapter, the conduction mechanism of oxide after quasi-breakdown (QB) is studied and modeled. It has been shown that post-QB stages can be separated into two distinct stages – recoverable and unrecoverable stages. In this chapter, further understanding of the mechanism of QB is sought, using thermal and bias annealing responses. While bias annealing studies in SILC and complete breakdown are extensive [5.1],[5.2], annealing behavior of post-QB oxides has not been well studied. Xu *et al*. have shown that annealing at 250^{o}C reduces interface traps significantly but with little reduction in SILC leakage current. Instead, reduction in SILC leakage current at higher temperature annealing (400^{o}C) is well correlated to oxide bulk traps reduction [5.3]. In the case of post-QB oxides, both interface traps and QB leakage current appear to reduce simultaneously at low temperature anneal (250^{o}C). While the study gives crucial insights to the annealing behavior of post-QB oxides, it is difficult to conclude on the roles of bulk and interface traps in post-QB oxides based on the evidences presented so far.

By using bias and thermal annealing of post-QB oxides together with trap measurements, this study attempts to clarify the relationships of the various traps to QB and its underlying mechanism. It was observed from reverse bias annealing experiment at room temperature that oxide bulk traps instead of interface traps, bear a closer correspondence to QB leakage current. As discussed in Chapter 3, post-QB can be distinguished into two phases – recoverable and unrecoverable QB. This is also observed using thermal annealing studies with disparate responses observed for the two stages. At the oxide thickness regime of 45 Å, it was observed that hole trapping predominates in the wearout stage. The studies provide strong supporting evidences that trapped holes contribute significantly to QB in thin oxides. At the same time, a

leakage path extending through the whole oxide thickness at initial QB is questionable since only reverse bias annealing reduces the QB leakage current, while the same polarity bias anneal further degrades it.

The organization of this chapter is as follows. Section 5.2 describes the device structures and experimental setup while Section 5.3 characterizes the annealing behavior of post-QB oxides under thermal and electrical bias anneal. Section 5.4 discusses the annealing mechanism in post-QB and proposes a hole trapping model with the formation of a localized trap region (LTR) to explain the various phenomena observed so far. Using this model, it is shown that both thermal anneal and bias anneal will result in a reduction of the LTR leading to electric recovery. However, the oxide traps induced under electric stress are not truly annealed out and recovery is only temporary. Finally, Section 5.5 concludes with a summary of the main experimental findings and conduction mechanism for QB.

5.2 Device and Experimental Setup

The devices and experimental setup for DCIV measurements are similar to previous chapter 3 (refer to section 3.1.1). The devices used in this experiment consist of p-MOSFET with channel area from 5-10 μm^2 and gate oxide thickness of 45 Å. All carrier separation measurements were carried under inversion conditions so that there are very little majority carrier resulting in almost negligible recombination within the space charge region. Under such conditions in a p-MOSFET, the source current I_S measures the hole current while the substrate or well current I_W measures the electron current. In subsequent experiments, the drain is not connected and left open as in the setup of previous chapter. For electrical stressing, Fowler-Nordheim (F-N) constant current stressing was performed with stress current density ranging from 20 mA/cm^2 to 100 mA/cm^2 on p-MOSFETs of varying channel area. Fig. 5.1 shows the experimental setup for the electrical bias annealing experiment. Carrier separation under inversion conditions [5.7] is performed at constant interval while negative gate polarity stressing using F-N carrier injection at constant current density was used to electrically stress the oxide till QB. At onset of QB, gate polarity is reversed for

reverse bias annealing. All measurements are carried out using the HP 4155B parameter analyzer.

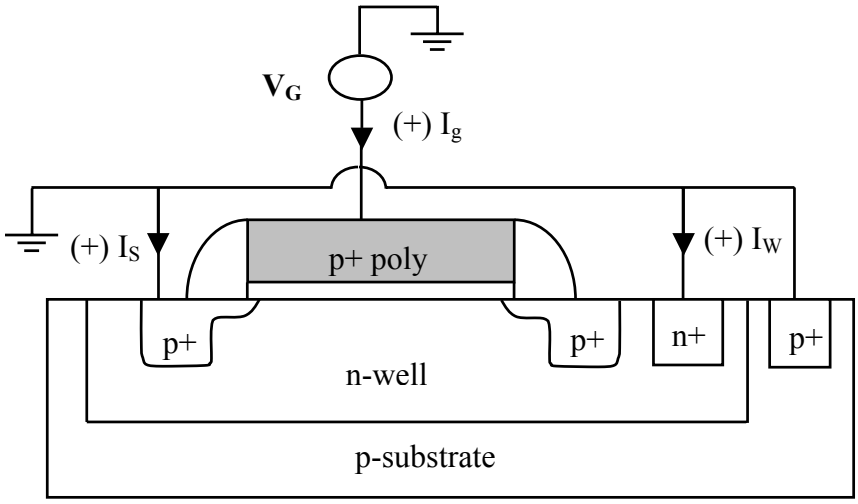


Fig. 5.1 Schematic illustration of carrier separation measurement setup for electrical bias annealing experiment.

To decouple the effect of electrical and thermal annealing, bias annealing is carried out at room temperature using the above setup while thermal annealing is carried out at elevated temperature. In the case of bias annealing, continual same polarity stress after onset of QB is considered as one form of bias annealing. In contrast, reverse bias annealing, refers to the application of a reverse gate bias stress to that of the original stress bias.

Post-QB thermal annealing is also studied by subjecting various samples at onset of QB or post-QB to high temperature annealing ranging from 150^{o}C to 350^{o}C for a period of 10 mins. Isothermal annealing was carried out in a N_2 ambient in a conventional horizontal furnace under high nitrogen (N_2) flow without electrical bias.

5.3 Characteristics of Electrical Recovery under Bias and Thermal Annealing

This section examines the effect of electrical bias and thermal anneal on post-QB oxides. Both DCIV and carrier separation measurements are carried out at successive log-time interval to monitor the traps evolution and leakage conduction mechanism respectively.

5.3.1 Bias Annealing of post-QB oxides

Figure 5.2 shows the evolution of the gate voltage after QB under continual constant current stress (CCS). CCS is used instead of CVS to limit the thermal effect due to current runaway at the onset of QB. A switching behavior can be observed with multiple stable voltages after QB. This is similar to that reported by Miranda *et al.* who have ascribed this behavior to the modulation of multiple conducting spots [5.8],[5.9]. It can be observed that gate voltage constantly fluctuates with momentary recovery to higher gate voltage although no full electrical recovery is observed.

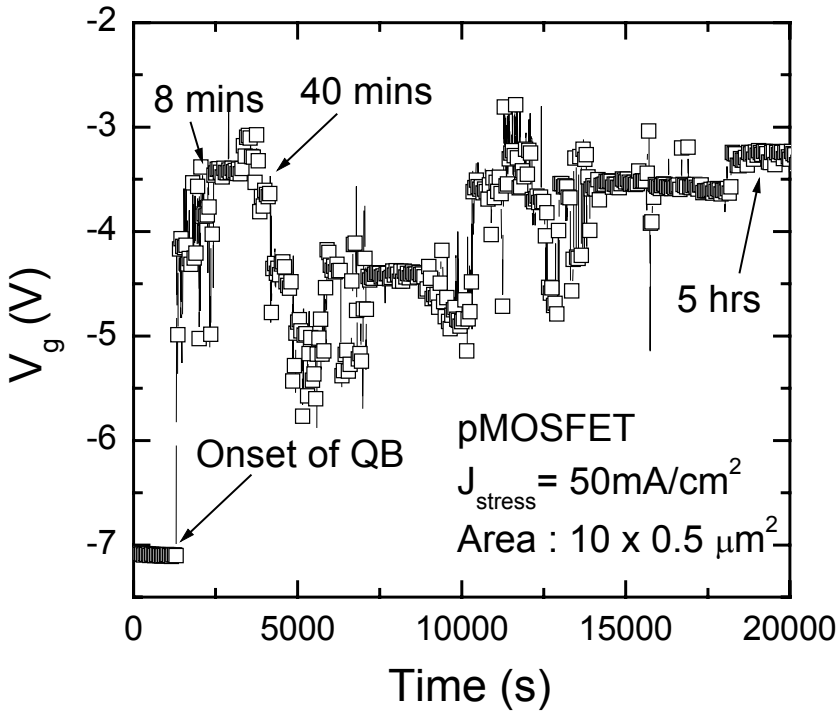


Fig. 5.2 Evolution of gate voltage under constant current stress till QB and post –QB positive bias annealing. (p-MSOFET, T_{ox} = 45 Å, W/L = 10/0.5 μm, J_{stress} = 50 mA/cm^2)

Continuous current stressing after onset of QB also shows that there are no changes in both the interface and bulk traps as shown in Fig. 5.3 and Fig. 5.4. The DCIV spectra overlaps after QB and this may be attributed to current channeling after QB. Positive bias annealing as shown by the continual gate current stress has momentary recovery in gate voltage but cannot recover the QB leakage current to pre-QB level. This will be contrasted with negative or reverse bias annealing whereby partial to almost full electrical recovery is observed.

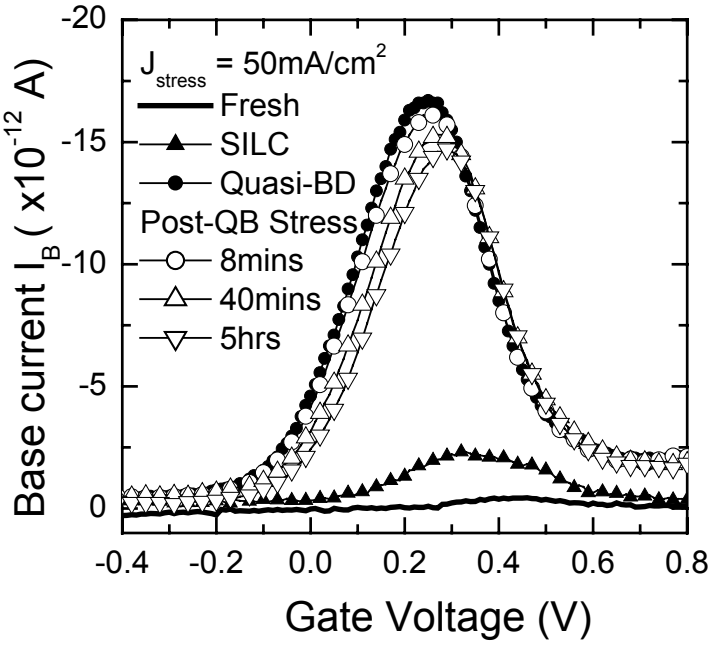


Fig. 5.3 DCIV spectra of p-MOSFET for Fig. 5.2, with stressing till QB and post-QB. After onset of QB, it can be observed that the recombination current, I_B spectra overlaps with subsequent decrease in the peak amplitude of I_B accompanied by a slight shift of V_{GB} for peak I_B to the right. (p-MSOFET, T_{ox} = 45 Å, W/L = 10/0.5 µm, J_{stress} = 50 mA/cm^2)

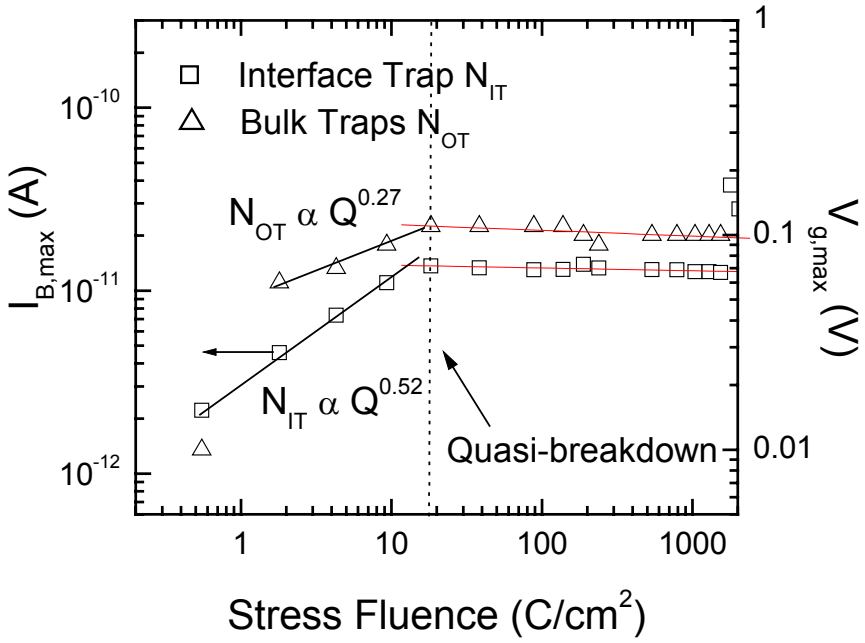


Fig. 5.4 Quantitative DCIV spectra measurement showing $I_{B,max}$ and ΔV_{GB} versus injected fluencies. Bulk and interface trap after QB show no further increment. (p-MOSFET, T_{ox} = 45 Å, W/L = 10/0.7 µm, J_{stress} = 50 mA/cm^2)

5.3.2 Reverse Bias Annealing after QB

The effect of reverse bias annealing on oxides stressed to QB using charge fluency of 142 C/cm^2 is shown in Fig. 5.5. Oxide was stressed till QB using CCS with current density of -100 mA/cm^2 (gate injection). At onset of QB, a reverse bias CCS of density 5 mA/cm^2 was immediately applied. Under reverse bias CCS, it can be observed that gate voltage shows initial fluctuation which are symptoms of QB. However, it recovered very quickly to a stable gate voltage equal to its pre-stressed voltage.

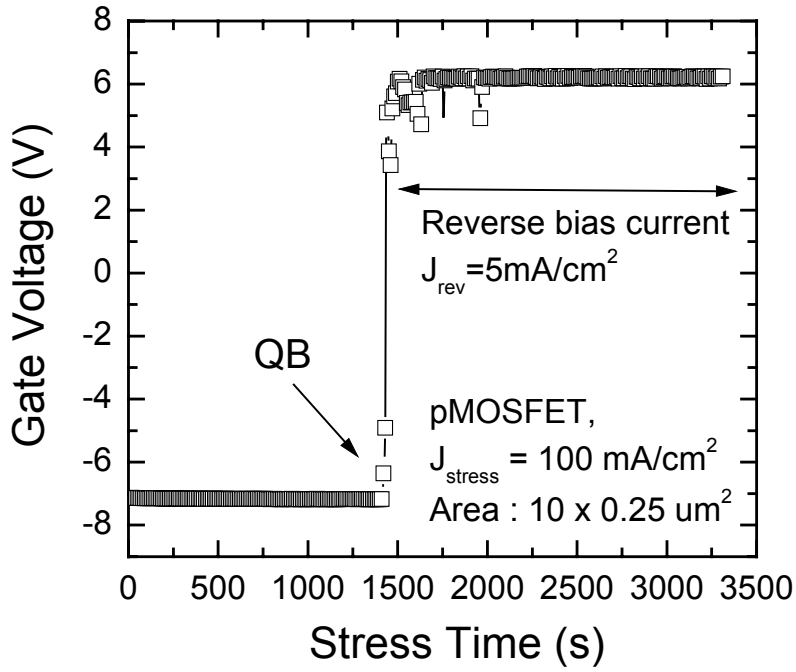


Fig. 5.5 Evolution of gate voltage under constant current stressing. At QB, reverse bias current J_{rev} = 5 mA/cm^2 is applied. (pMOSFET, Tox = 45 Å, W/L = 10/0.25 μm, J_{stress} = 100 mA/cm^2)

Figure 5.6 shows the post-QB I-V characteristics as well as those after various reverse bias annealing time. The reduction of post-QB gate leakage current at high gate voltage increases with the annealing time, till it matches the fresh state when additional 2.5 C/cm^2 of reverse bias fluency is applied. This behavior is quite different from the positive bias annealing and shows that QB states can be annealed using reverse electric bias. This electric annealing effect is independent of the stressing polarity and the same effect is also observed when a positive gate bias stress is applied, followed by a reverse negative gate bias anneal.

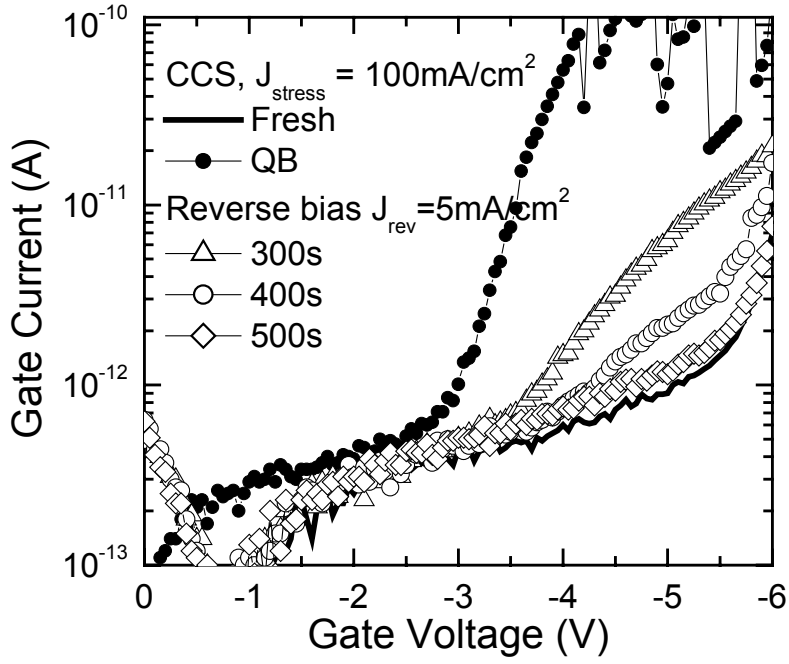


Fig. 5.6 Associated I-V characteristics of oxide at various stages of electrical stressing as shown inset. Reverse bias anneal applied after onset of QB. It can be seen that gate leakage recovers back to fresh after 500 s of reverse bias anneal. Sample used is the same as Fig. 5.5.

In order to further understand the evolution of trap generation during bias annealing, DCIV measurements are taken prior to and after reverse bias anneal. Fig. 5.7 shows the DCIV spectra of p-MOSFET stressed to QB using constant current stressing. QB was attained at fluency of 70 C/cm^2. Further additional positive bias stressing results in insignificant changes to interface and bulk traps as shown by the overlapping of DCIV spectra. Upon application of a reverse bias, DCIV spectra shift positively by about 0.1 V, showing a decrease in oxide bulk traps while the magnitude of the base recombination current I_B remains almost constant. Since the peak of I_B is proportional to interface traps, it can be inferred that reverse bias reduces the oxide bulk traps but not the interface traps. It can also be observed that a short application of the reverse bias accounts for a significant positive shift in the gate voltage for the DCIV spectra. Subsequent reverse bias annealing beyond 15 s results in very minor voltage shift till almost the same measured value at initial fresh state.

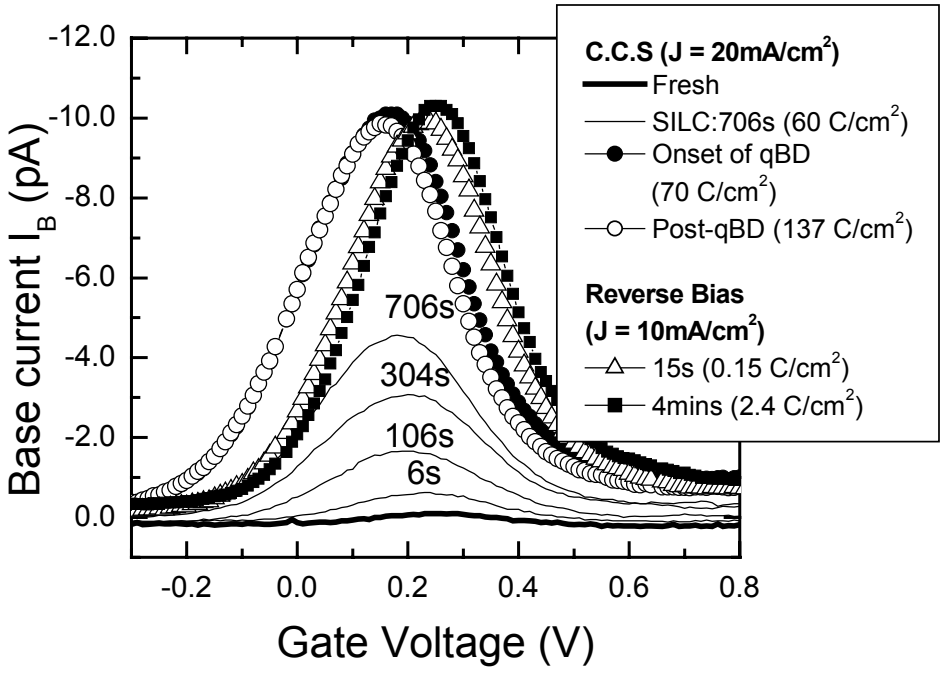


Fig. 5.7 DCIV spectra for p-MOSFET oxide at various stages of electrical stressing. The sample used is the same as Fig. 5.8.

Figures 5.8 shows the associated I-V characteristics for the samples monitored for its traps using DCIV in Fig. 5.7. It can be seen that post-QB gate leakage reduces significantly after reverse bias annealing for 15 s. A direct correspondence of reduction in bulk traps at reverse bias anneals and post-QB gate leakage current reduction under reverse bias annealing was observed. Fig. 5.9 shows the change in I_B (which reflects the interface traps) and voltage shift of peak I_B (which reflects the oxide bulk traps) at onset of QB and post-QB with application of reverse bias anneal. It can be observed that with the application of a reverse bias anneal, gate leakage current recovers to SILC states (Refer to Fig. 5.8), together with the recovery of oxide bulk traps to almost its fresh level. The results suggest that QB leakage current can be attributed to oxide bulk traps instead of interface traps as first suggested by Guan *et al*. [5.10]. The discrepancies may be attributable to the methods of measuring the traps. In the case of [5.10], an indirect measurement of oxide traps using the $\Delta J_g/J_g$ which tends to probe the region closer to the injecting cathode, may have introduced the discrepancies.

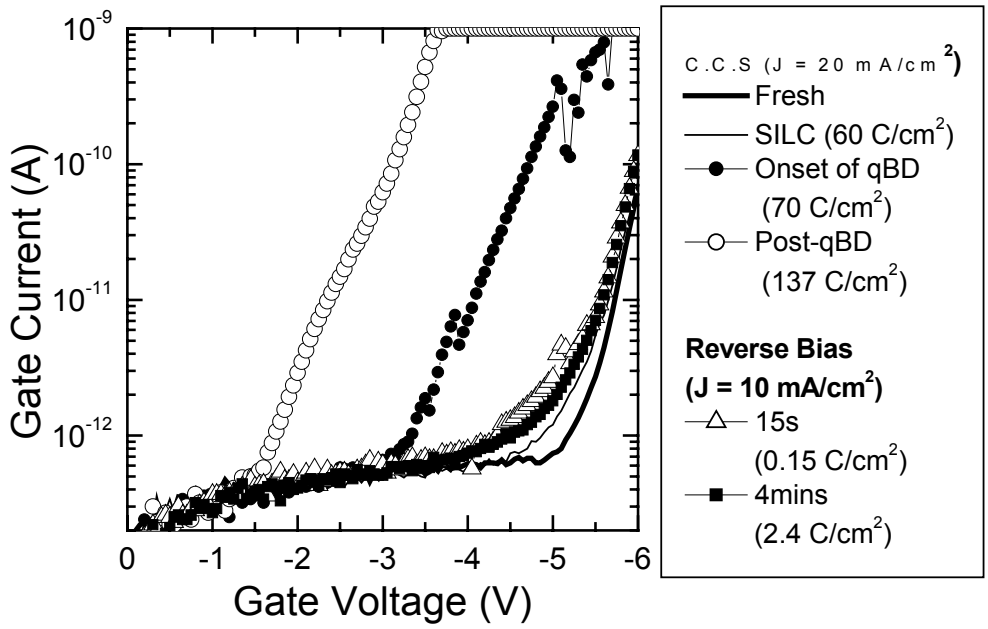


Fig. 5.8 Associated I-V characteristics at various stages of constant current stressing. After quasi-breakdown is attained, same polarity stressing is continued before application of a reverse bias stress. Gate leakage current after application of reverse bias shows reduction till SILC level. (T_{ox} = 45 Å, W/L = 10/0.5 μm, p-MOSFET).

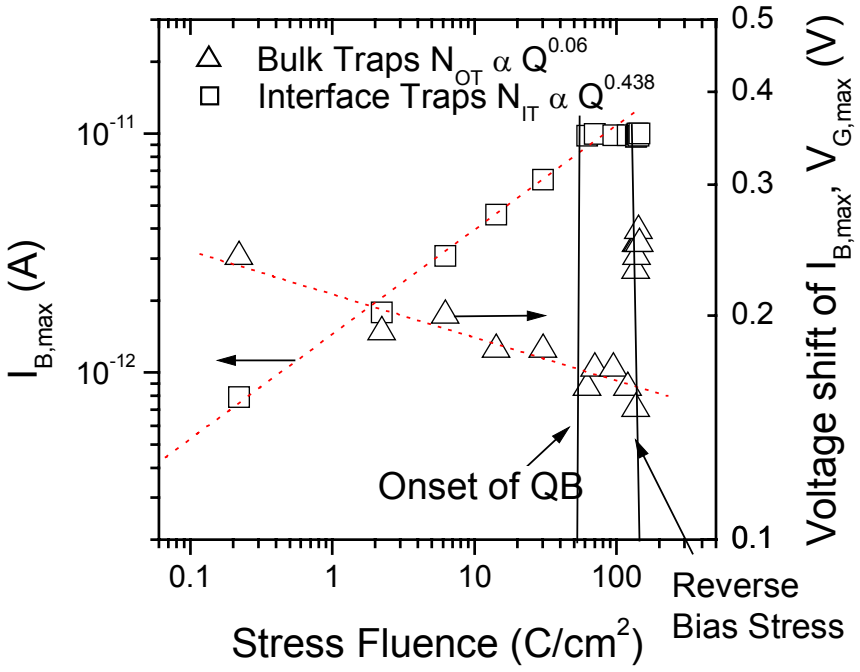


Fig. 5.9 Quantitative DCIV spectra measurement at various stages of current stressing. Peak recombination current I_B is related to interface traps while the lateral shift of peak I_B is related to oxide bulk traps. At reverse bias anneal, oxide bulk traps recover to initial values at fresh state while interface traps remains unchanged.

Figure 5.10 shows the successive cycle of bipolar stressing and bias anneal prior to and after QB. Before QB, both positive and negative gate polarity stressing results in a relatively stable gate voltage under constant current stressing. Post-QB oxides however show large fluctuation in the gate voltage for both positive and negative gate polarity stressing. Moreover, it was observed that bias annealing did not truly anneal out the bulk traps but merely deactivates it with rebound of QB observed, as shown by the large leakage current (as indicated by the low gate voltage) observed at the next successive reverse bias stressing due to reactivation of the oxide bulk traps.

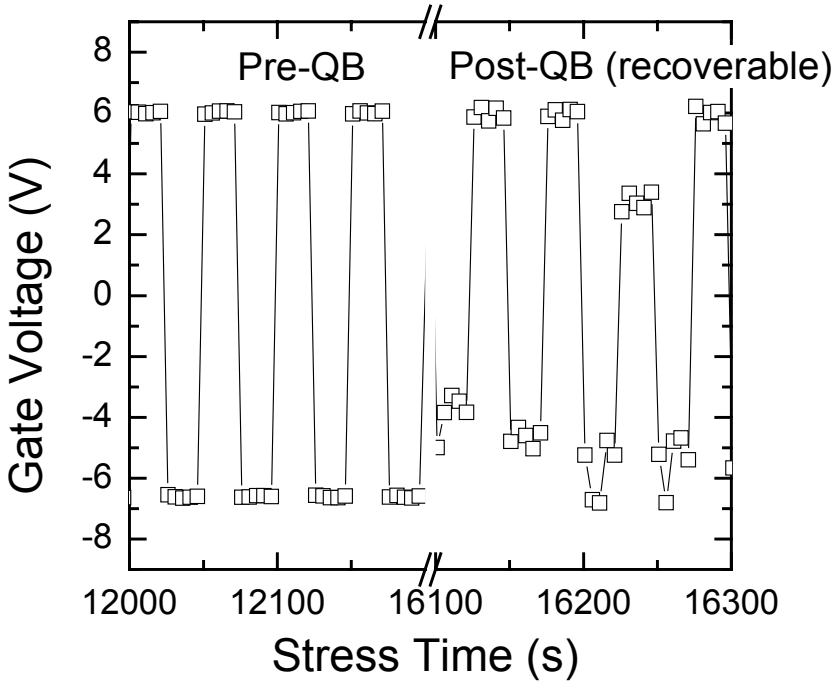


Fig. 5.10 Variation of gate voltage under bipolar constant current density J_{stress} = +/- 10 mA/cm^2 at pre-QB and post-QB with successive alternating stressing and bias annealing. (T_{ox} = 45Å, W/L = 10/0.2 μm, n-MOSFET)

5.3.3 Thermal Annealing after QB

In Chapter 4, it has been shown that there exist two stages within QB characterized by their electrical recoverability. As such, the thermal annealing experiment will be conducted on post-QB oxides at both the electrical recoverable and unrecoverable stages. Fig. 5.11 shows the DCIV spectra of oxide stressed to QB (recoverable) and subjected to successive different temperature annealing.

Recoverable QB stage is attained by stressing with a low current density and by stopping electrical stress immediately upon QB. It can be observed that increasing temperature anneal results in lowering of both oxide bulk and interface traps as shown by the progressive left shift and lowering of peak amplitude of the base recombination current I_B respectively. From the DCIV spectra, it can be observed that the locus of the recombination current I_B maxima, is shifted positively upon application of successive thermal annealing. Irrespective of the annealing temperature, the locus of the post-anneal samples follows a similar locus as that of pre-QB oxide but with a positive shift of about 0.05 V. The results suggest that there is a fast transient component of either positive trap de-trapping or negative charge generation under thermal annealing for the temperature range used here: 150^{o}C to 350^{o}C. Since there is no electrical bias during the thermal annealing, additional negative charge generation seems unlikely, hence suggesting that positive charge inherent in post-QB oxides de-trap easily under a thermal anneal above 150^{o}C. Unlike bias annealing, no significant changes in the gate leakage current was observed for thermal annealing till 250^{o}C as

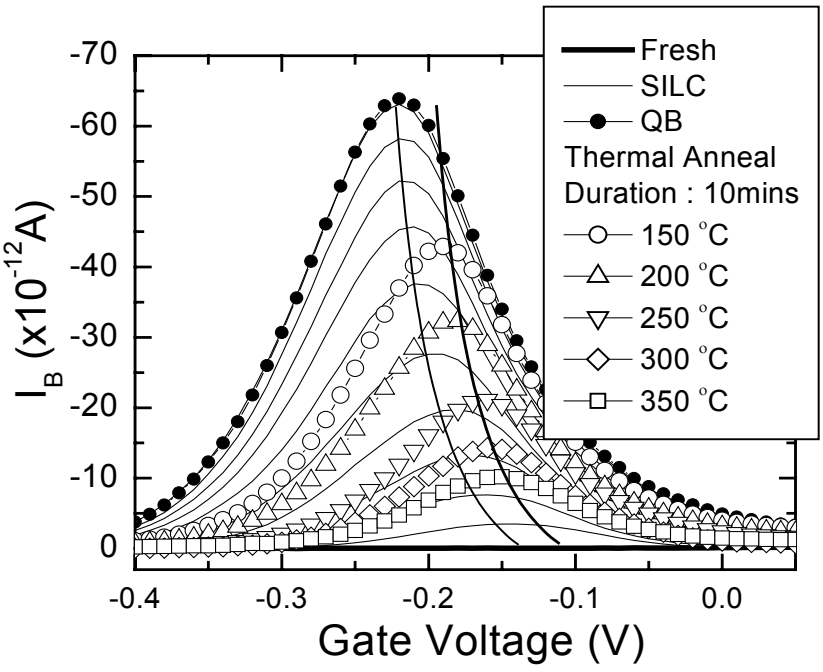


Fig. 5.11 DCIV spectra of p-MOSFET stressed to QB. Post-QB thermal annealing performed at various temperatures as shown inset. The thin line linking up the maxima of the I_B spectra reflects the level of oxide bulk traps during SILC while the thick line reflect bulk trap level due to the thermal annealing. It can be observed that thermal anneal results in both a positive shift in the spectra and reduction in the maxima of I_B showing reduction in interface traps and bulk traps.(T_{ox} = 45 Å, J_{stress} = 50 mA/cm^2, W/L = 10/1.0 μm, p-MOSFET).

shown in Fig. 5.12. Beyond 250°C, gate leakage current is further enhanced resulting in complete breakdown after thermal annealing at 400°C. Between 250°C to 350°C, it can be observed that gate leakage current first increases at low field and then subsequently reduces. This will be further discussed in section 5.4.1. The disparate response between thermal annealing and electrical bias annealing is puzzling. In the case of thermal annealing, it can be observed that annealing at 400°C results in a right shift of the DCIV spectra to almost the pre-stress level. While this reduces the gate leakage current in the case of electrical reverse bias anneal, gate leakage current did not decrease under thermal annealing conditions. The experimental data as shown in Fig. 5.12 does not agree with the data for the first stage of QB as reported by Ang *et al*. [5.11] who showed that there exists two distinct stages in QB, but fits the characteristics of the annealing behavior for the second stage of QB. One possibility for the difference in behavior although both samples are stressed till onset of QB and subjected immediately to thermal annealing, is the instability inherent in the initial stage of QB.

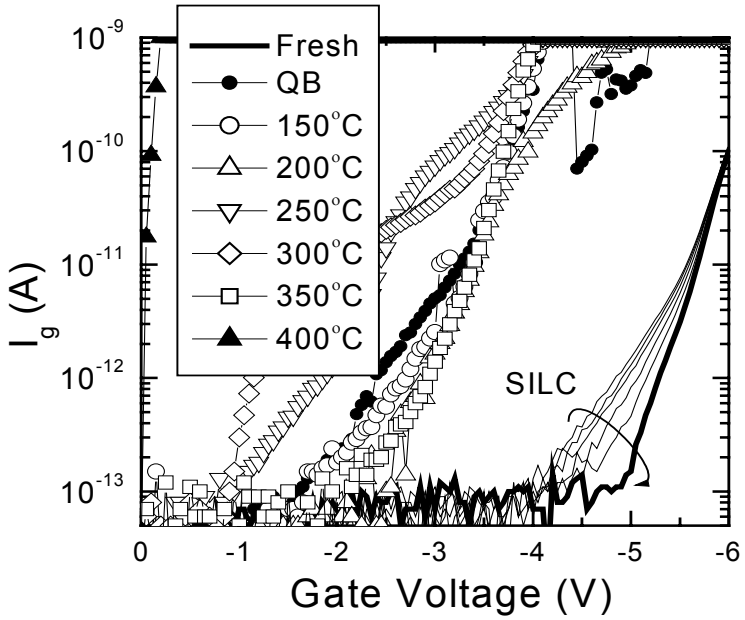


Fig. 5.12 Gate current leakage current after QB and with post-QB thermal annealing at successively higher temperature for 10mins each. (T_{ox} = 45 Å, J_{stress} = 50 mA/cm^2, W/L = 10/1.0 μm, p-MOSFET).

Figure 5.13 shows the DCIV spectra for p-MOSFET stressed to QB and subjected to a constant thermal annealing temperature of 200°C for increasing

duration of annealing time. As compared to Fig. 5.11, it can be observed that the effect of constant temperature annealing results mainly in bulk traps reduction as shown by the right shift of the DCIV spectra while interface traps are reduced marginally. Even then, the voltage shift of the DCIV spectra is still significantly less than that for higher temperature. The result shows trap annealing at elevated temperature has higher dependencies on annealing temperature rather than annealing time. Despite the insignificant changes in trap density, significant effect is observed for gate leakage current.

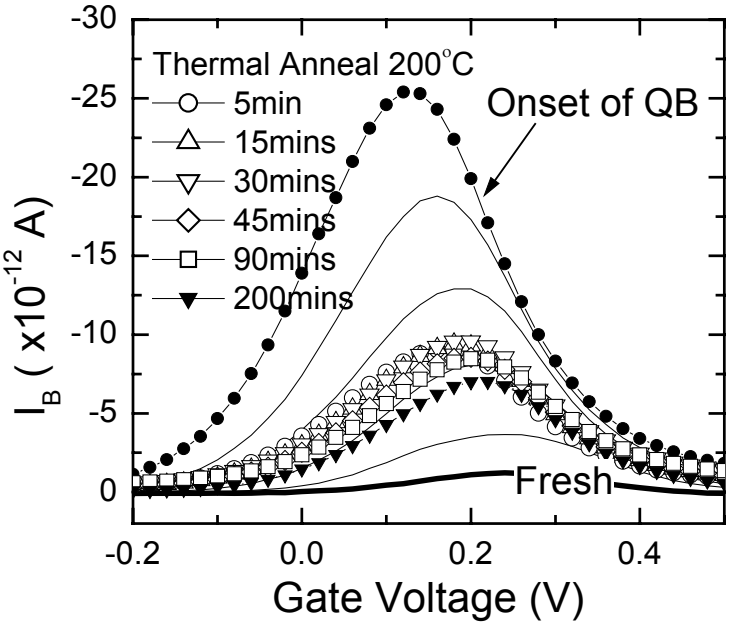


Fig. 5.13 DCIV spectra of p-MOSFET stressed to QB. Post-QB thermal annealing carried out at 200°C for varying period of annealing durations from 5 mins to 200 mins. (T_{ox} = 45 Å, J_{stress} = 20 mA/cm^2, W/L = 10/1.0 μm, p-MOSFET).

Figures 5.14 – 5.16 shows the carrier separation measurements taken after various annealing time at low anneal temperature of 200°C. Fig. 5.14 shows the gate leakage current reduction at onset of QB and after thermal annealing at 200°C. Unlike the reduction in post-QB gate leakage to its fresh level for electrical bias anneal, thermal anneal can only reduce post-QB leakage current by a certain amount, without full recovery observed. This result is consistent with experimental result of Okandan *et al.* [5.12] who showed that QB cannot be completely annealed out even at 400°C. It is further observed that this final gate leakage current has a profile parallel to the F-N leakage current at fresh state.

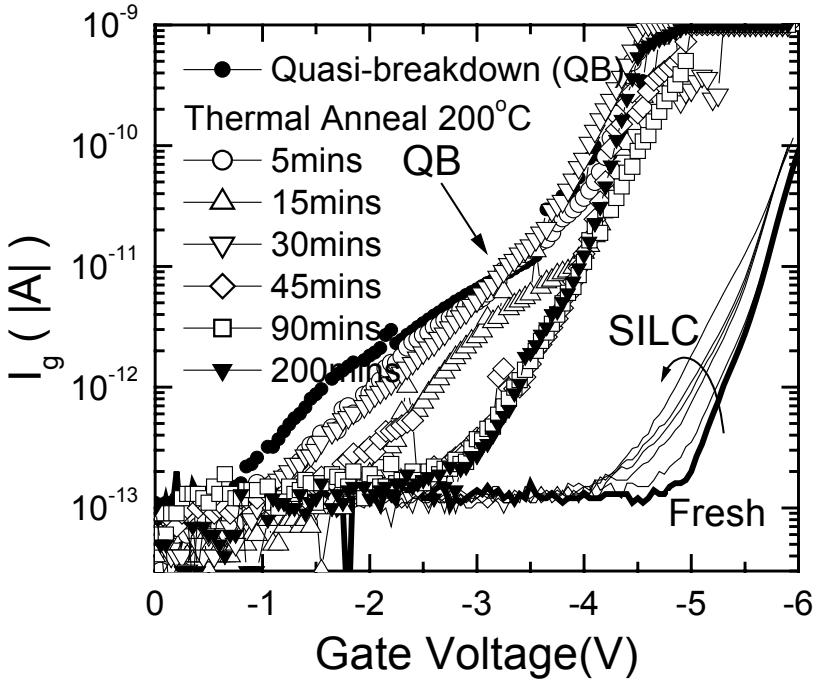


Fig. 5.14 Gate leakage current at QB and after post-QB thermal annealing treatment at 200°C for varying period of time as shown inset.

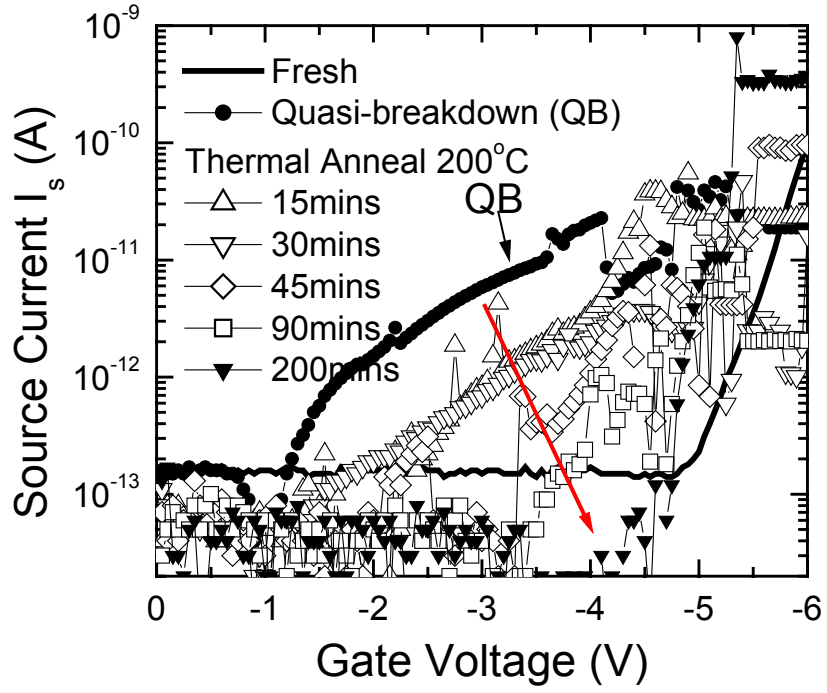


Fig. 5.15 Carrier separation measurement of source current component (holes current) for p-MOSFET under successive thermal anneal. Sample used is the same as Fig. 5.14

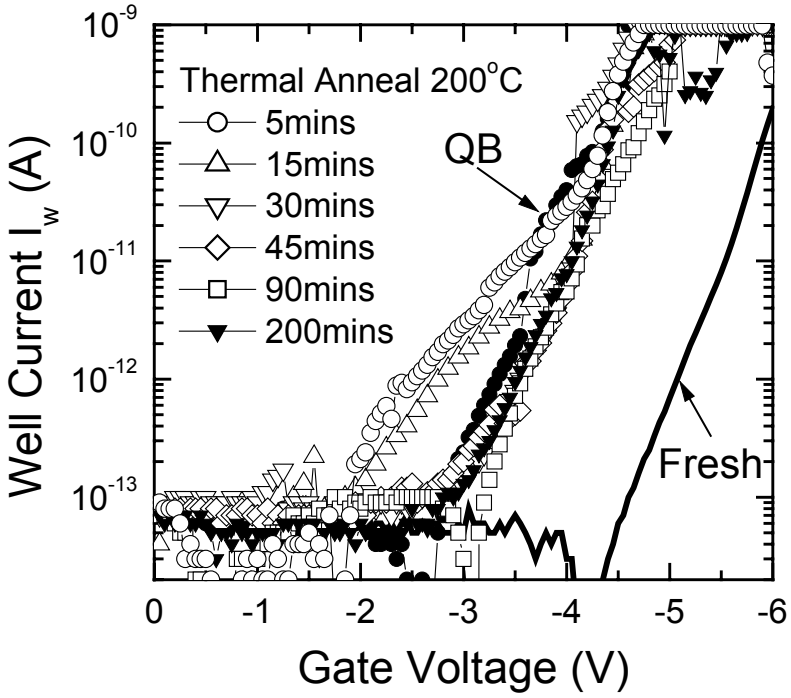


Fig. 5.16 Carrier separation measurement of well current component (electron current) for p-MOSFET under inversion conditions after successive thermal anneal. Sample used is the same as Fig. 5.14

From Fig. 5.15, it can be observed that hole current (as shown by source current) reduces continuously under increasing annealing duration. With 200 mins of thermal annealing at 200°C, source current has virtually been reduced to fresh leakage level. Electron current, on the other hand, as shown by well/substrate current as shown in Fig. 5.16 also reduces continuously with increasing thermal annealing time but saturates at a certain level after 45 mins of annealing and does not reduce any further even after further annealing till 200 mins. The combined effect of electron and hole current is observed in the gate leakage current as shown in Fig. 5.14. With progressive annealing at 200°C, gate leakage current changes from a hole dominated leakage current to an electron dominated leakage current, with a final saturation level as determined by the electron leakage current.

5.3.4 Recoverable and Unrecoverable QB states

All previous thermal annealing experiments were conducted on oxide stressed to onset of QB and hence are still in the electrical recoverable stage. In the subsequent experiment, the thermal response of oxides stressed to unrecoverable QB is observed.

Fig. 5.17 - Fig. 5.19 shows the carrier measurement for gate oxides stressed to unrecoverable QB and subjected to increasing duration of thermal anneal at 200°C.

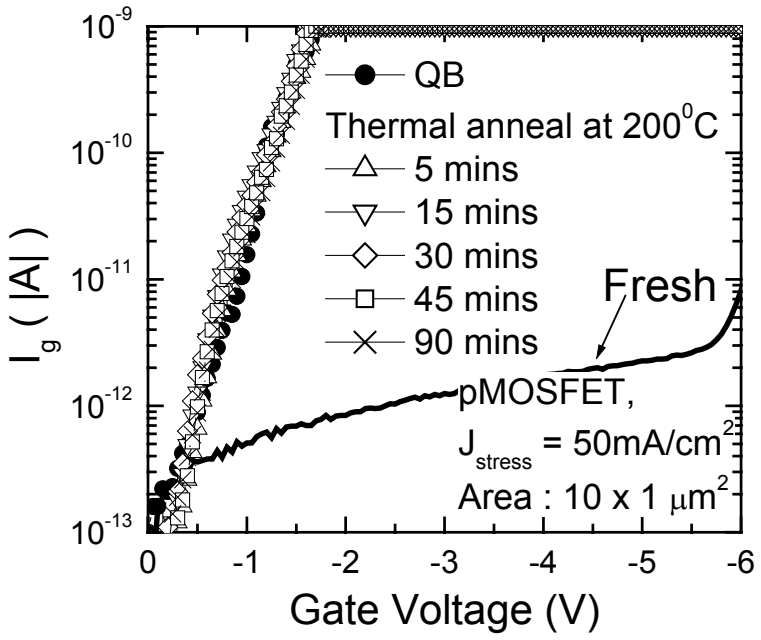


Fig. 5.17 Gate I-V characteristics for p-MSOFET stressed till unrecoverable QB with post-QB thermal annealing at 200°C for varying period of durations as shown inset. (T_{ox} = 45 Å, J_{stress} = 50 mA/cm^2, W/L = 10/1 μm, p-MOSFET)

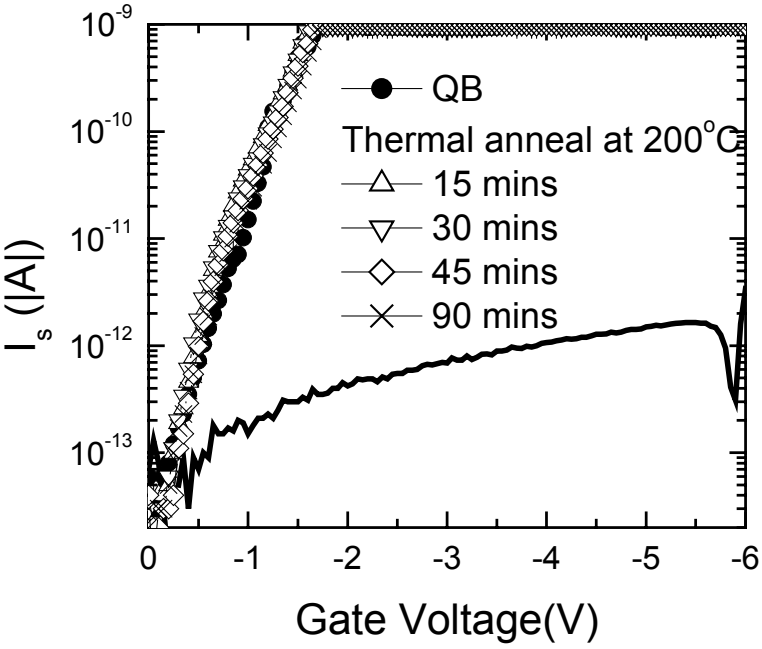


Fig. 5.18 Carrier separation for source I-V characteristics on the same p-MSOFET used in Fig. 5.17 with post-QB thermal annealing treatment.

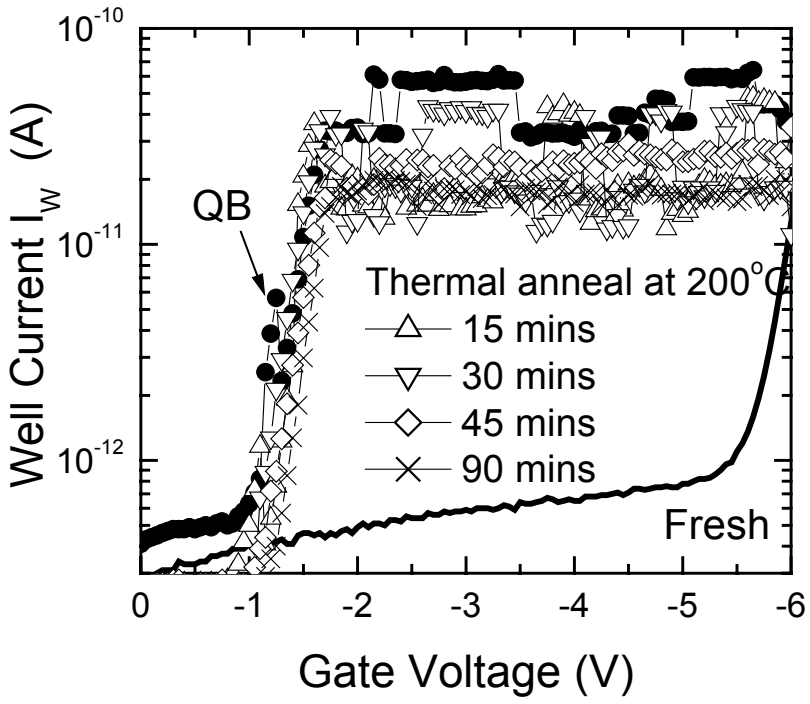


Fig. 5.19 Carrier separation measurement showing substrate I-V characteristics for p-MOSFET used in Fig. 5.17 with post-QB thermal annealing treatment.

It can be observed that thermal annealing has no effect on the gate leakage current at unrecoverable QB. This response corresponds to that of electrical bias annealing and shows that unrecoverable QB is a distinct stage of it own. Moreover, the leakage current at unrecoverable QB is very stable unlike recoverable QB with leakage current much higher than that of recoverable QB at both low and high field but still significantly smaller than complete breakdown.

Carrier separation measurements also demonstrate an interesting result. As can be observed from Fig. 5.18, the magnitude of hole leakage current is significant at low field bias from 0 V while that of electron current (as shown by I_w in Fig. 5.19) is negligible below $V_g < 1V$, which corresponds exactly to the Si bandgap barrier above which, valence band electron from p+gate can directly tunnel to substrate.

5.3.5 Combined Annealing Results: Bias and Thermal Anneal

It is intuitive from the previous discussion and argument that an electric bias coupled with thermal anneal may be able to anneal out most of the trapped holes. Fig. 5.20 shows the I-V characteristics of an oxide subjected to CCS till QB and

subsequently to a bias and thermal annealing. It can be seen that the bias anneal recover most of the gate leakage current to its pre-QB level while thermal anneal results in minor but further reduction in gate leakage till almost fresh, unstressed level, consistent with the previous results for separate bias and thermal annealing. Fig. 5.21 shows the associated DCIV spectra for both bias and thermal anneal performed sequentially. Similar results, as before, such as positive shift in DCIV voltage showing bulk traps reduction, charge compensation or charge de- trapping due to bias anneal and interface traps reduction under thermal anneal are observed. However, even with bias and thermal annealing, the 'recovered' oxide goes to QB relatively easily showing that no true recovery has occurred as shown in Fig. 5.22. The results are almost similar to a post-QB oxide but subjected to reverse bias annealing and show that thermal annealing performed under such conditions is unable to truly anneal out the underlying traps causing the initial QB.

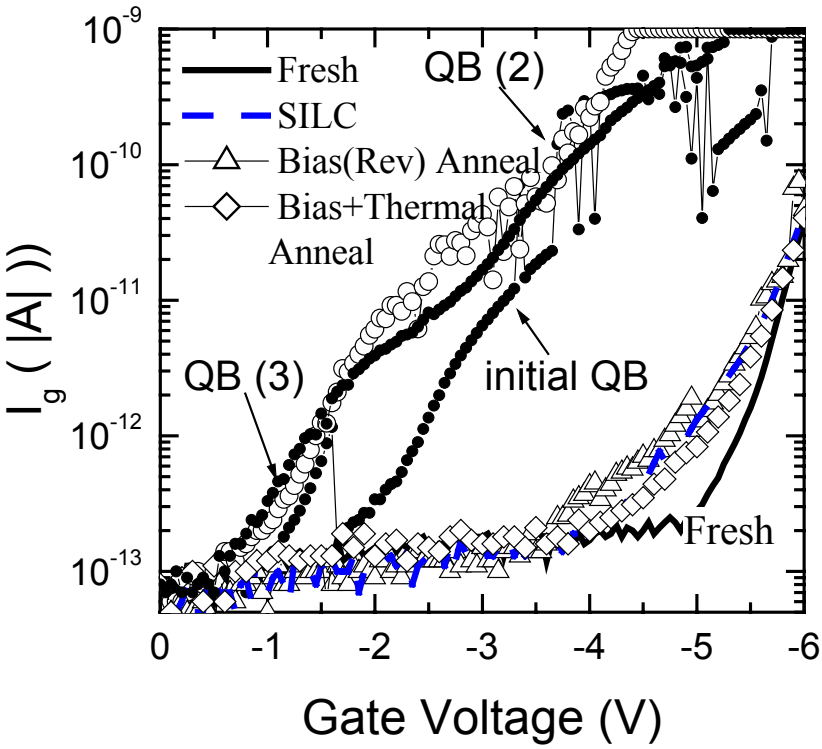


Fig. 5.20 Gate leakage current at fresh stage and after electrical stresses. Oxide was stressed till onset of QB as shown by initial QB and continual stressing result in QB(2) before being subjected to reverse bias and thermal anneal. The 'recovered' oxide was then subjected to additional electrical stress till second QB as shown by QB(3).

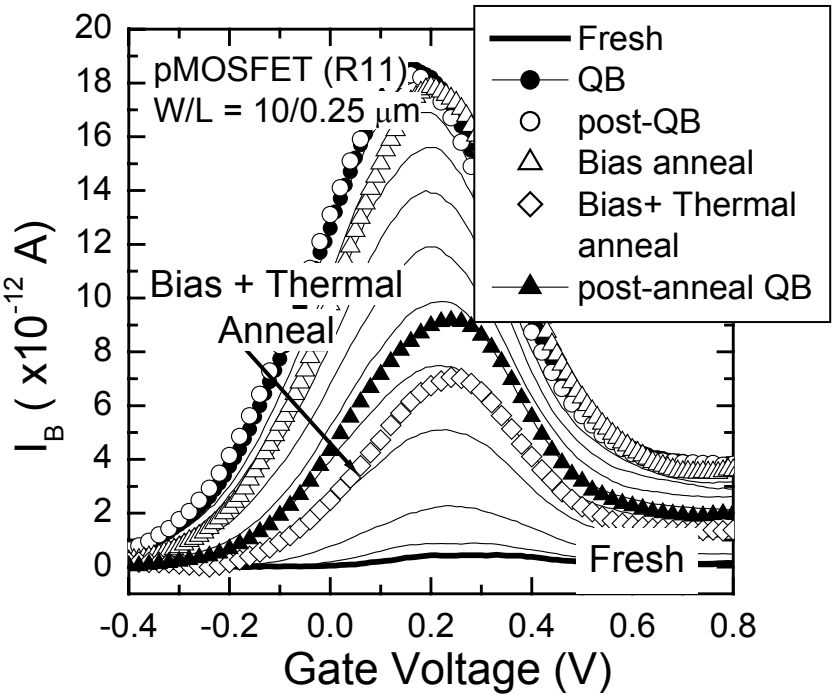


Fig. 5.21 Associated DCIV spectra of oxide stressed till QB and subjected to bias and thermal anneal. It can be observed that the combined effects of bias and thermal anneal results mainly in a positive shift in the DCIV spectra of the post-QB oxide and reduces I_B to a lower level respectively.

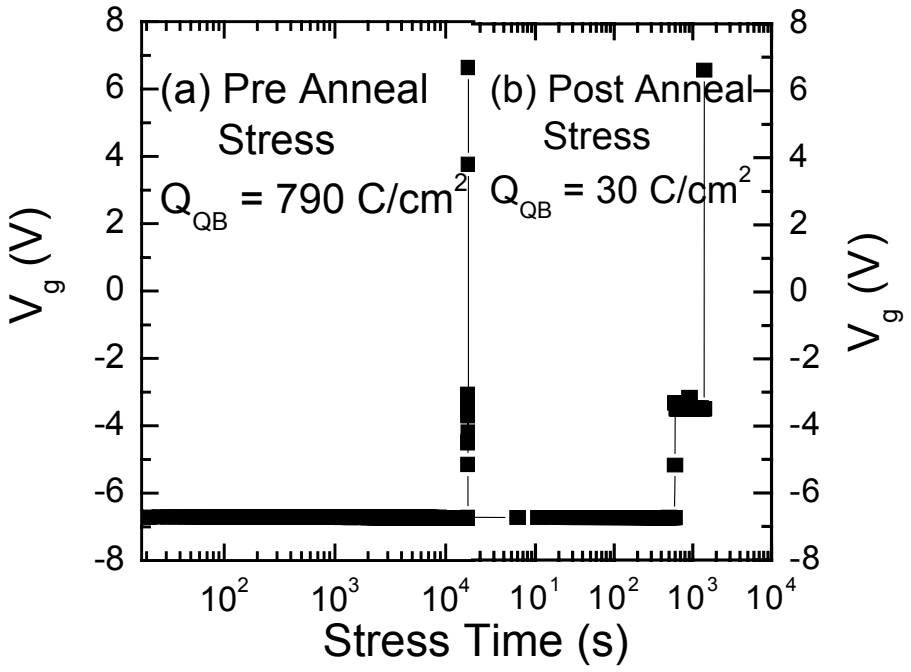


Fig. 5.22 Evolution of the gate voltage of oxide subjected to CCS till QB as shown in (a) and subsequently subjected to bias and thermal anneal. (b) shows the subsequent evolution of the gate voltage when the 'recovered' oxide is re-subjected to CCS till a second QB.

5.4 Discussions for bias and thermal annealing

In the previous section 5.3.2, Fig. 5.9 shows that it is bulk traps that lead to QB leakage current based on reverse bias anneal experiment. Negative voltage shift of DCIV spectra shows that the responsible oxide bulk traps have a positive charge, showing that hole trapping is predominant. Based on bipolar current stressing, it was observed that the reverse bias annealing does not annihilate the defect responsible for QB but merely deactivate them. The leakage current reduction under post-QB reverse bias annealing can then be explained by the following possible mechanisms. (1) Neutral Electron Traps (NETs) compensation due to occupancy of injected electrons which reduces the tunneling sites. (2) Trapped holes annealing due either to electron-hole compensation as a result of electron tunneling from substrate/gate or thermal detrapping of holes. While both mechanism are possible, it is clear that mechanism (1) does not cause any true annihilation of defects and cannot explain why significant electron trapping only occurs under reverse bias. Mechanism (2) appears more likely and correlates with the initial observation that hole trapping results in QB leakage current. The mechanism of hole trapping has been well studied and hole compensation due to electron trapping is evidenced from various annealing studies [5.13].

Figure 5.23 shows the microscopic model for hole trapping due to strained Si-Si bond resulting from a missing oxygen precursor. The association of hole trapping at the E' center is well evidenced by electron-spin-resonance (ESR) studies. Electrons tunneling from silicon can restore the net electrical neutrality thus nullifying the electrical influence of the trapped holes without actually removing the trapped holes. Using this model, it was observed that a consistent explanation can be provided for the experimental data presented.

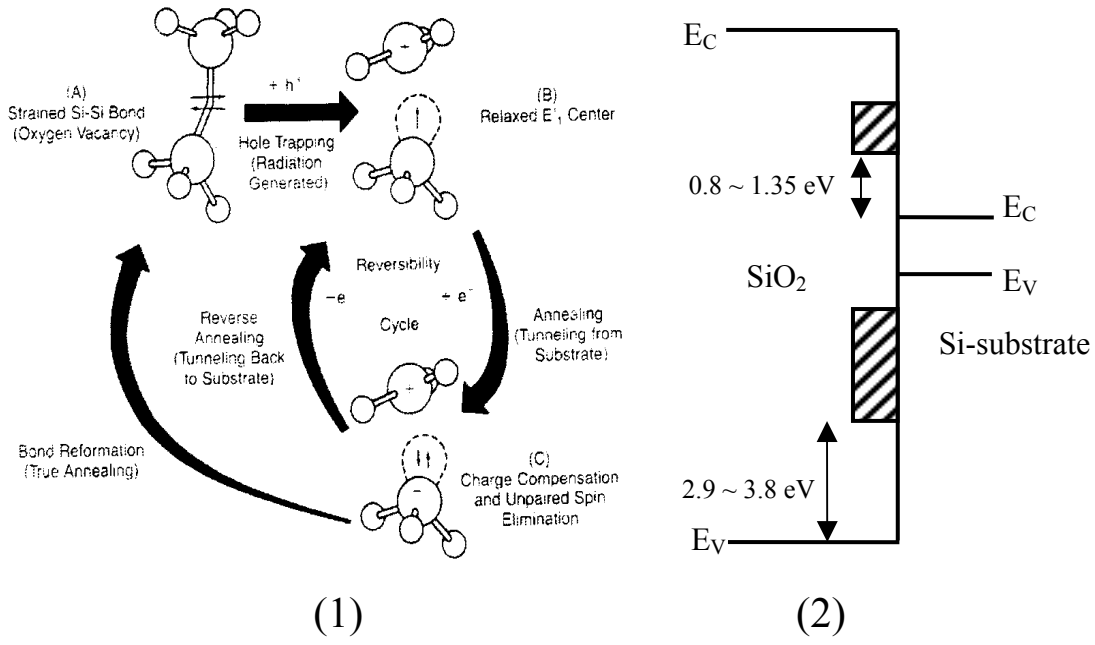


Fig. 5.23 (1) Microscopic model of hole trapping (A)-(B) forming an E' center and detrapping (C) along with charge compensation and bond reformation as proposed by Lelis et al. [5.13] (2) Two spatial equivalent trap levels that electrons can tunnel to, corresponding to the ground and excited state of the E' center. (After [5.13])

5.4.1 Mechanism of thermal and bias anneal

It was observed in Fig. 5.12 that thermal anneal did not result in significant reduction in gate leakage current as compared to bias annealing. In addition, constant temperature annealing at 200^{o}C as shown in Fig. 5.14 shows a modest decrease in gate leakage after 45 mins of annealing but saturates thereafter without further decrease. Carrier separation measurement shows that limiting factors can be attributed to electron leakage current which cannot be annealed out. The result suggests that hole traps with energy level below the Si valence band can be easily annealed out, while deep hole traps with energy above the Si conduction band cannot be readily anneal out. Considering the energy level of trapped holes as shown in Fig. 5.23 (2), the phenomenon could be easily understood considering that valence band electron from the Si substrate (n-well) can easily tunnel through the oxide and neutralize the E' center for trap level below the Si valence band as shown in Fig. 5.24. On the other hand, for electrons to reach the trap level above the Si conduction band, it would require additional thermal energy, with compensation process being limited by the annealing temperature. In the case of bias anneal, this barrier is lowered by the

electric field allowing electrons to tunnel from the Si conduction band. As a result, hole leakage current which depends on hole traps with energy level below the Si

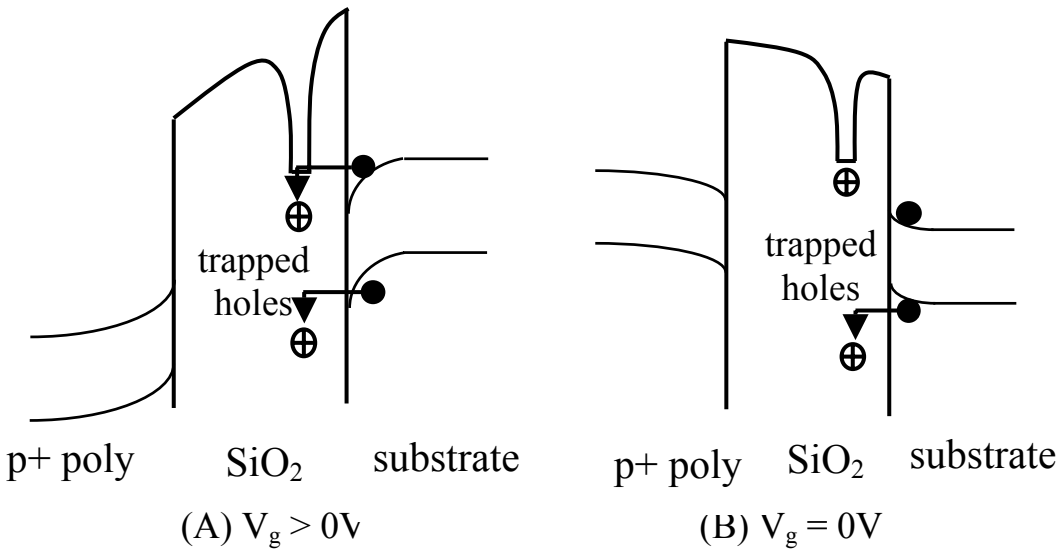


Fig. 5.24 Schematic diagram illustrating (A) Reverse bias annealing (B) Thermal annealing without bias. With bias anneal, both levels of trapped holes can be annealed while in thermal annealing, hole trap with energy level above Si conduction band requires electrons with energy above conduction band to be deactivated.

valence band is easily reduced under low temperature anneal (200^{o}C) as shown in Fig. 5.15 while that of electron leakage current remains significant, limited by the thermal annealing temperature.

The mechanism for the annealing of trapped holes resulting in the reduction of post-QB stress-induced leakage current is likely to be explained by two possible mechanisms, namely: (1) lattice relaxation resulting in true bond reformation and (2) a metastable neutral center where electron and hole are associated but do not recombine. Leslis *et al*. have proposed a microscopic model for hole trapping associated with a positively charged E' center [5.13]. From his high temperature bias annealing experiment, it was observed that a significant fraction of trapped holes that are apparently annealed out could be re-activated under a reverse polarity. This has been ascribed to localized holes (E'γ centers) which form a metastable, dipolar complex without restoring the Si-Si bond upon electron capture [5.13],[5.14]. In this experiment, since the annealing temperature is relatively low, it is believed that thermal annealing is due to electron compensation at the metastable neutral center

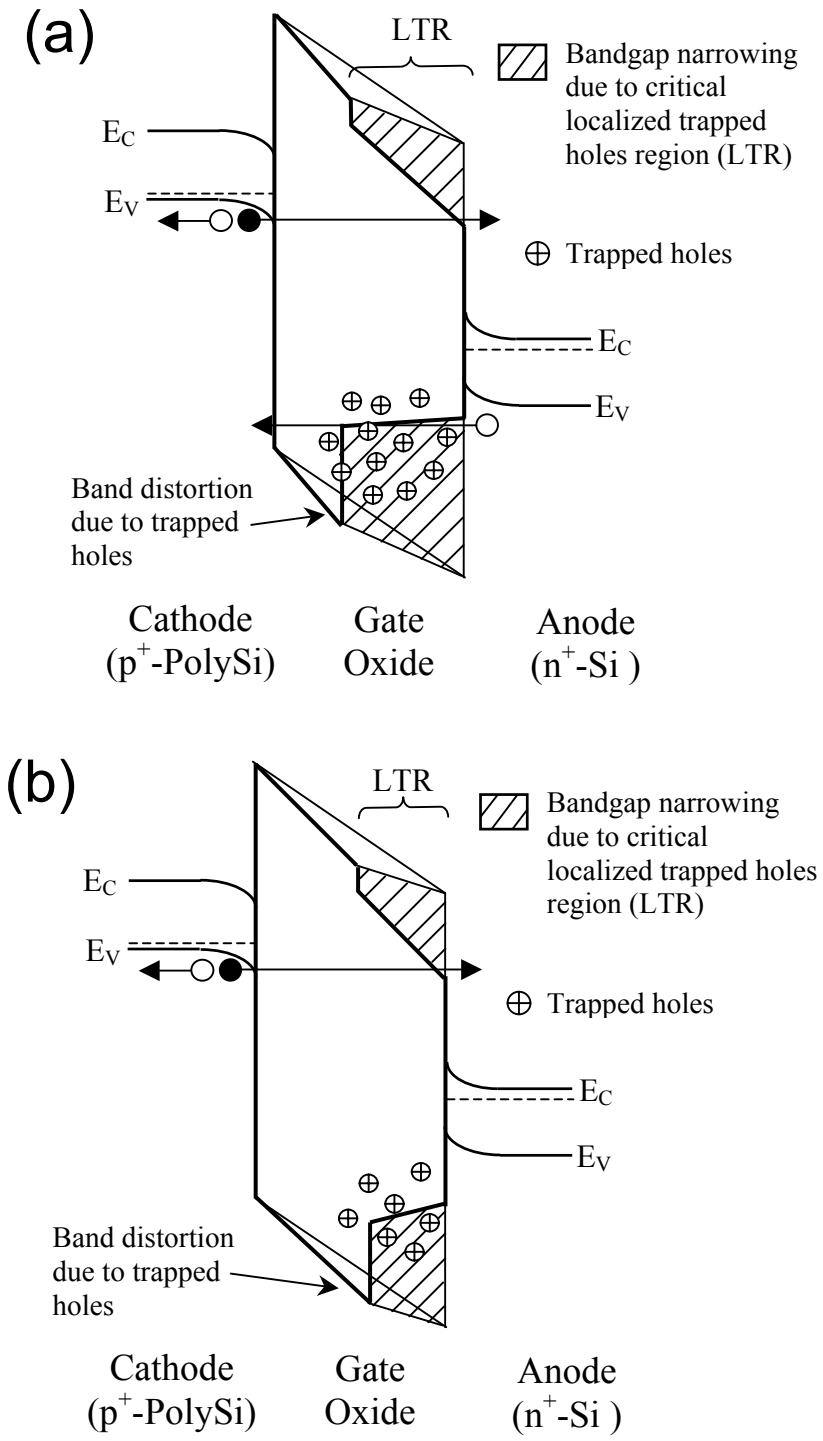


Fig. 5.25 (a) Proposed localized trap region (LTR) model: hole trapping predominantly at anode causes an energy band distortion and results in F-N conduction for electrons and direct tunneling for holes at the recoverable QB. (b) Thermal annealing results in electron-hole pair compensation and the reduction of localized trap region (LTR). The shaded regions represent bandgap narrowing due to formation of LTR while the thick lines show the resultant oxide energy band.

rather than due to lattice relaxation. During thermal annealing, electrons are thermally excited from the anode and are injected into the oxide thus compensating some of the trapped holes. As a result, the localized trap region (LTR) shrinks continuously till eventually, hole current cannot directly tunnel to the cathode. Shallow trap levels near to the anode also cannot be easily detrapped or compensated with electrons by thermal annealing alone, thus leaving behind a finite amount of trapped holes and certain regions of LTR as shown in Fig. 5.25 (b). This accounts for the $QB_{threshold}$ current observed after thermal annealing. Using the above model, the anomaly observed in Fig. 5.12 for the disparate response of gate leakage to annealing temperature can be explained using similar analogy to the model proposed by Xu *et al.* [5.3]. In his model, it was proposed that there exist of two competitive components: a interface damage region and shortening path within the bulk oxide. As annealing temperature is increased, the interface damage component shrinks but the shortening path increases, resulting in an eventual increase in gate leakage. In our case, thermal annealing will lead to electron compensation of trapped holes close to the anode resulting in a quenching of the LTR. In addition, experiments by other researchers have shown that hole trapping and de-trapping is a reversible process [5.15],[5.16]. Due to thermal energy, holes can then detrap and be re-trapped in an adjacent site, leading to an increase in the lateral dimension of the LTR. In the case of Fig. 5.12, gate leakage at low field increases initially when subjected to thermal anneal of 250^{o}C to 300^{o}C due to faster lateral expansion of existing trapped holes spatial location via detrapping/re-trapping process at high temperature. Above 350^{o}C, diffusion rate increase but is more than offset by the higher electron compensation from both the anode and cathode due to higher density of injected electron and its higher energy state. As a result, the effective LTR may be shrunken, leaving only uncompensated shallow traps near to the anode with gate leakage similar to that at onset of QB as shown in Fig. 5.12. The competition between the detrapping/compensation of the trapped holes close to anode and diffusion of the trapped holes leads to a disparate gate leakage response to different temperature annealing.

In contrast, it is speculated that the unrecoverable QB is achieved when the LTR spans the entire thickness of the oxide. In this case, there will be no high potential barriers for electron and hole between cathode and anode, and the injected carriers can freely flow across the gate oxide. Under this situation, both trapping and detrapping can occur easily. This will hinder the net electrons compensating the trapped holes, and therefore the recovery of the oxide. At this stage, electrical recovery is not achieved since any injected electrons can freely flow to anode without compensating the trapped holes.

5.5 Summary

Using thermal and electrical bias annealing experiment, it has been observed that post-QB oxides can be recovered using a reverse bias annealing. The recovery is only temporary and subsequent stressing quickly revert the oxide back to its QB state. At low temperature, annealing is able to reduce the holes current till pre-QB level while electron current saturates after a certain level and does not reduce any further. The result suggests that electron and holes conduction proceed by different traps unlike SILC. The response of recoverable and unrecoverable QB stage is also contrasted. Under both electrical and thermal annealing, it was observed that unrecoverable QB remains very stable without any recovery observed. The disparate response shows that both are distinct stages within QB.

Reference

[5.1] T. Endoh, K. Shimizu, H. Iizuka, and F. Masuoka, "A new write/erase method for the reduction of stress-induced leakage current based on the deactivation of step tunneling sites for flash memories," *IEDM Tech. Dig*. pp. 49-52, 1994.

[5.2] A. Meinertzhagen, C. Petit, M, Jourdain, and F. Mondon, "Stress-induced leakage current reduction by a low field of opposite polarity to the stress field," *J. Appl. Phys*., vol. 84, pp. 5070-5079, 1998.

[5.3] Z. Xu, B. J. Cho and M. F. Li, " Annealing behaviour of gate oxide leakage current after quasi-breakdown", *Microelectronics Reliability*, vol. 40, no.8-10, pp. 1341, 2000.

[5.4] J. Sune and E. Miranda, "Post Soft Breakdown Conduction in SiO_2 Gate Oxides," *IEDM Tech. Dig.*, pp. 533-536, 2000.

[5.5] A. Shanware, J. P. Shiely, and H. Z. Massoud, "Extraction of Gate Oxide thickness of n- and p-channel MOSFETs below 20 Å from the substrate current resulting from valence-band electron tunneling," *IEDM Tech. Dig*., pp. 815-818, 1999.

[5.6] S. I. Takagi, N. Yasuda, and A. Toriumi, "Experimental Evidence of inelastic Tunneling and new I-V model for stress-induced leakage current," *IEDM Tech. Dig.,* pp. 323-326, 1996.

[5.7] H. Guan, B. J. Cho, M. F. Li, Y. D. He, Z. Xu, and Z. Dong, "A study of quasi-breakdown mechanism in ultra-thin gate oxide by using DCIV technique," *IEEE*, *7th Proc. IPFA*, pp. 81-84, 1999.

[5.8] E. Miranda, J. Sune, R. Rodriguez, M. Nafria, and X. Aymerich, "Soft breakdown fluctuation events in ultrathin SIO_2 layers," *Appl. Phys. Lett.,* vol. 37, pp. 490-492, 1998.

[5.9] E. Miranda, J. Sune, R. Rodriguez, M. Nafria, X. Aymerich, L. Fonseca, and F. Campabadal, "Soft breakdown conduction in ultrathin (3-5nm) gate dielectrics," *IEEE Trans. Electron Dev*., vol. 47, no. 1, pp. 82-88, 2000.

[5.10] H. Guan, B. J. Cho, M. F. Li, Z. Xu, Y. D. He, and Z. Dong, "Experimental evidence of interface-controlled mechanism of quasi-breakdown in ultra-thin gate oxide," *IEEE Trans. on Electron Dev.,* vol. 48, No. 5, p. 1010, May 2001..

[5.11] C. H. Ang, "Ultra-thin Gate Oxide Reliability," *NUS Ph.D thesis,* pp. 168-169, 2000.

[5.12] M. Okanadan, S. J. Fonash, B. Maiti, H. H. Tseng, and P. Tobin, "Wearout, quasi-breakdown, and Annealing of ultrathin dielectrics : impact on complementary metal oxide semiconductor performance and reliability," *Electrochemical Solid-state Lett*., vol. 11, pp. 583-584, 1999.

[5.13] A. J. Lelis, T. R. Oldham, H. E. Boesch Jr., and F. B. Mclean, "The nature of the trapped hole annealing process," *IEEE Trans. on Nuclear Science,* vol. 36, no. 6, pp. 1808-1815, 1989.

[5.14] S. P. Karna, A. C. Pineda, R. D. Pugh, W. M. Shedd, and T. R. Oldham, **"**Electronic structure theory and mechanisms of the oxide trapped hole annealing process," *IEEE Trans. Nucl. Sci.,* vol. 47, pp. 2316-2321, 2000.

[5.15] J. F. Zhang, I. S. Al-kofahi, and G. Groeseneken, "Behavior of hot hole stressed SiO_2/Si interface at elevated temperature," *J. Appl. Phys*., vol. 83, pp. 843-850, 1998.

[5.16] H. S. Witham, and P. M. Lenahan, "Nature of the *E*' deep hole trap in metal-oxide-semiconductor oxides," *Appl. Phys. Lett.,* vol. 51, pp. 1007-1009, 1987.

Chapter 6

Degradation and Breakdown Mechanism in Ultra-thin Oxides

6.1 Introduction

In the previous chapters, a new mechanism relating to thin oxides was studied. Quasi-breakdown (QB) [6.1]-[6.3], Soft breakdown (SBD) [6.4] or Mode B SILC [6.5] which was observed predominantly in thin oxide below 50 Å was described and its underlying mechanism was investigated and modeled. The emergence of new degradation mechanism is a consequence of the continual thinning of the conventional silicon dioxides used for gate dielectric application. Due to this aggressive scaling, many aspects of breakdown mechanism in SiO_2 may have changed. First and foremost, the conduction mechanism in ultra-thin silicon dioxide is changed as its operating voltage condition is progressively scaled down. As SiO_2 is scaled below 30 Å and for low gate voltage below 3.2 V, gate leakage through the dielectric is governed mainly by direct tunneling (DT) through a triangular barrier rather than by the conventional Fowler Nordheim (FN) tunneling. Unlike FN tunneling, DT tunneling is highly dictated by the oxide thickness, and as a rough approximation, the gate leakage current increases by one order of magnitude for every 2 Å decrease in oxide thickness. Due to this change in conduction mechanism, many other mechanisms may also have changed.

In thicker oxides (> 35 Å), gate leakage is due primarily to stress-induced leakage current (SILC) which has both a transient and steady-state component [6.6],[6.7]. This has been primarily ascribed to positive charge-assisted tunneling [6.8], localized spots/filamentary path with lower barrier height [6.9], trap-assisted tunneling (TAT) through neutral traps [6.6],[6.10] and electron-hole recombination at positively charged centers [6.7]. Beside SILC, another phenomenon observed for oxide thinner than 50 Å is quasi-breakdown (QB) as described in the previous chapter. Based on electrical recovery of QB for 45 Å oxides, it has been shown that

QB leakage current can be correlated to the positive oxide trapped charges instead of interface traps, showing that QB leakage path is partially due to trapped holes in the oxide [6.3]. On the other hand, it has also been shown that QB is triggered by a critical number of interface traps. These inconsistencies can be reconciled by the localized trap region model at the anode comprising of trapped holes due to anode hole injection.

In sub-4 nm thick oxides, however, positive charge buildup almost completely disappears due to direct tunneling of any trapped charges. Recently, Nicollian *et al*. and Ghetti *et al*. have shown that low voltage SILC (LVSILC) in oxide less than 35 Å is dominated by interfacial trap tunneling mechanism [6.12],[6.13]. Using temperature dependence studies, an anomalous increase in gate leakage at low voltage close to flatband condition is observed, and this was found to have weak temperature dependencies [6.13]. The explanation for this phenomenon is attributed mainly to electron tunneling through interface states [6.13],[6.14]. On the other hand, using hot-carrier stressing and temperature studies, Meinertzhagen *et al.* have observed that not all interfacial traps can be measured by LVSILC and the level of interface states determined by LVSILC is not correlated with eventual oxide breakdown [6.14].

In oxides thicker than 30 Å, dielectric breakdown can be clearly differentiated into quasi-breakdown (QB) [6.1]-[6.4] or conventional breakdown depending on the severity of the degradation [6.4],[6.17]. In both cases, the result is immediate device failure due to the significantly high gate leakage current and thermal effects accompanying the breakdown occurrence. As oxide thickness is scaled downwards, it is predicted that conventional complete breakdown will be less likely to occur due to lower power dissipation [6.18]. In ultra-thin gate oxides, with thickness T_{ox} < 14 Å, it is hard to observe breakdown or quasi-breakdown under typical stress conditions. In most cases, QB is expressed as onset of gate voltage or current fluctuation [6.16]. Moreover the gate leakage current shows a continuous increase over the entire period of electrical stress. This increase in gate direct tunneling current coupled with lower operating voltage tends to obscure the gate current increase due to QB occurrence. Weir *et al.* [6.19] have shown that a single occurrence of QB, especially in 25 Å dielectrics, may not degrade device switching performance. Due to the fact that the

device may not fail after a single occurrence of quasi-breakdown, the lifetime of ultra-thin oxide becomes highly dependent on the statistics of multiple quasi-breakdown occurrence. In this respect, very detailed studies [6.22]-[6.29] have been done and shown that subsequent QB statistics can be successfully modeled using multiple breakdown spots statistics. Consequently, Wu *et al*. [6.20] have also proposed a new failure criterion, using a dual voltage time-dependent dielectric wearout (TDDW) to characterize and monitor device failure for 18-27 Å oxide. For ultra-thin oxides, Monsieur *et al* [6.21] have further observed that current increase in 17-24 Å oxides is progressive and has characteristics dissimilar to QB. In their detailed study, it has been observed that different device areas have almost identical wear-out current, thus leading them to conclude that progressive breakdown (PBD) dynamics is independent of device area. It was suggested that the PBD are spatially correlated and grow from a single or a few degraded spots [6.21],[6.27]-[6.29]. It has also been observed using emission microscopy that PBD can be characterized by an increase in area at a single degraded spot [6.21]. On the other hand, evidences using multiple spots modeling have shown that there are no correlation between the degraded spots [6.22],[6.24]-[6.26] and enhanced gate leakage may be explained by independent multiple breakdown spots.

In this chapter, new experimental findings in the degradation and breakdown mechanism for ultra-thin silicon dioxides in the thickness regime from 14 Å ~ 20 Å are reported. Section 6.2 depicts the device fabrication and measurement setup for the various electrical measurements techniques used in this chapter. Section 6.3 describes the experimental results for thin 20 Å oxides while Section 6.4 describes the degradation in gate leakage current observed in 13 Å ultra-thin oxides. In Section 6.3, it is reported that gate leakage current in 20 Å oxides increases in a ‘steplike’ fashion and this is correlated with interfacial QB rather than the conventional bulk and interfacial induced QB as described in Chapter 4 for 45 Å oxides. Moreover, as the thickness of oxide shrinks further to less than 14 Å, the magnitude of the gate current density increases in ultra-thin gate oxide eventually becomes too high to be acceptable for normal device operation as shown in Section 6.4. A lifetime criterion based on the increase in gate leakage current is proposed as described in Section 6.4. Our study

shows that the area-dependence of the gate leakage current density increase in 13.4 Å oxides is different from that in thicker oxide films, indicating a localized and discrete property of the leakage current. It has also been observed that the oxide lifetime based on the new lifetime criterion is shorter when the gate area is smaller, as opposed to the conventional area dependence of time-to-breakdown test. Section 6.5 describes a simple model consisting of multiple degraded spots and it has been shown that localized gate leakage current in 13 Å oxides can be described by Weibull's statistics for multiple degraded spots. Finally Section 6.6 concludes with a summary of oxide degradation for ultra-thin and thin gate oxides in the thickness regime of 13 Å to 20 Å.

6.2 Device Fabrication and Experimental Setup

CMOS devices were fabricated using a standard dual-gate CMOS technology on (100) *p*-type and *n*-type substrate using 0.15 μm CMOS technology. Gate oxides with a physical thickness of 13 Å to 20 Å were grown using rapid thermal oxidation (RTO) at 800^{o}C and 950^{o}C respectively. In the case of the 13 Å gate oxides, a post-oxidation exposure to high-density nitrogen plasma was performed to incorporate nitrogen near the top of the gate dielectric for gate leakage current reduction and improve boron penetration susceptibility. Fig. 6.1 shows the measured C-V curves, which can be fitted very well with the simulated results for direct tunneling in 13 Å oxides with quantum mechanical corrections. Dual poly-silicon gate was used and the channel area of the samples used in this study ranges from 2 μm^2 to 2500 μm^2.

For constant current stress (CCS) and constant voltage stressing (CVS), both positive and negative gate polarity were used while source, drain and n-well were grounded. Direct-current current-voltage (DCIV) measurement using vertical bipolar transistor formed by p^+ source/n-well/*p*-substrate was used to measure interface and bulk traps. In our measurements, n-well and *p*-substrate are grounded. Drain, which is not used, in this case is left unconnected. Bias conditions for the forward biased vertical bipolar transistor were V_{EB} = + 0.3 V and V_{BC} = 0 V. Recombination current at the interface states in the channel region was monitored via the base recombination current I_B from source region of the vertical parasitic p/n/p bipolar transistor. From

the base current measurements, the interface trap and bulk trap generation can be measured via the peak value of I_B and the shift of gate voltage at peak I_B respectively. All measurements were carried out using a HP4155B parameter analyzer.

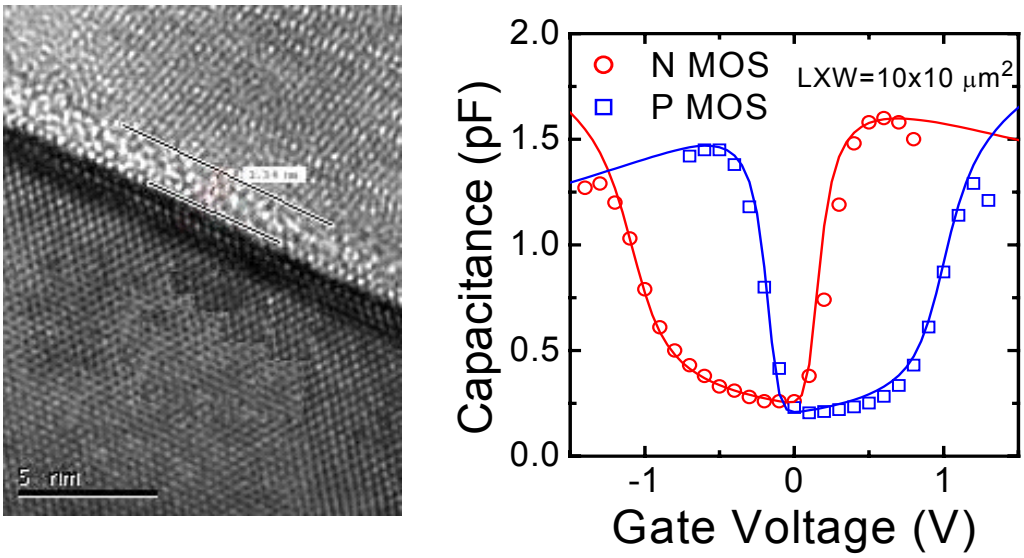


Fig. 6.1 HRTEM cross section of 13 Å gate oxide (left) and C-V measurements and simulation results (solid lines) fitting to 13 Å oxide thickness by Berkeley QMCV modeling (right).

6.3 Thin Oxide (20 Å) QB Degradation Characteristics

In the subsequent sections, the degradation observed in thin (20Å) is described. It is found that the characteristics of silicon dioxides in these thickness regimes, are entirely different from that of thicker oxide (> 45 Å) which has been described in the previous chapters 4 and 5.

6.3.1 Comparison between Thick and Thin oxide: Interface Trap Enhanced Tunneling (ITET)

Figure 6.2 shows the gate voltage under constant current stress (J_{stress} = -50 mA/cm^2) for both thick and thin oxide. In thicker oxide (T_{ox} = 45 Å), gate voltage V_g as shown by the solid symbols, initially increases showing negative charge trapping near the cathode. QB is observed in thick 45 Å oxides after 69 C/cm^2 of charge fluences. At onset of QB, a large decrease in the gate voltage can be observed with the change in gate voltage of around 30%. This orders of magnitude decrease in the gate

voltage under CCS at QB is reflected in a corresponding increase in gate leakage under a constant voltage. Due to this increase in gate leakage especially at low voltage, QB renders the device unsuitable for use in low power application. As a result, QB is often considered as an oxide failure for thick oxides in conventional lifetime prediction studies. In thinner oxide (T_{ox} = 20 Å) however, it is observed that discrete step decreases in gate voltage occurred after about 10^4 C/cm^2 of injected electron fluency with a comparatively smaller decrease in gate voltage than in thicker

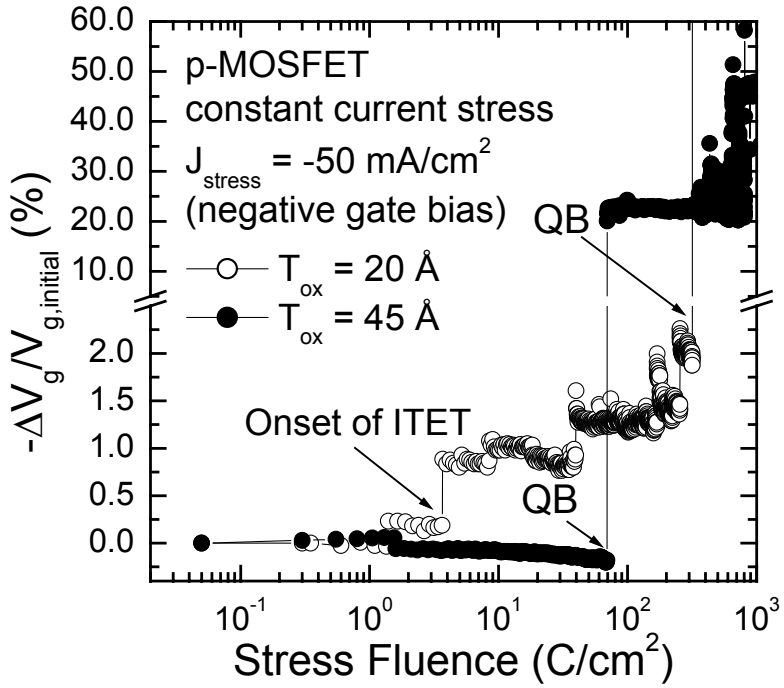


Fig. 6.2 Evolution of gate voltage under constant current stressing (J_{stress} = -50 mA/cm^2) with gate injection for thick (45 Å) and thin (20 Å) gate oxide. (T_{ox} = 45 Å & 20 Å, Channel Area = 10 μm^2, p-MOSFET).

oxide as shown in Fig. 6.2. This discrete step change in the gate voltage under constant current stress in thin oxide is very similar to QB observed in the thicker oxide. The key difference is in the magnitude of the gate voltage change, which may be explained by the higher direct tunneling current in the thin oxide. Besides magnitude differences, it is also observed that this step-like change in gate voltage for thin oxide is not completely similar to that of thicker oxide as will be shown subsequently. To differentiate this smaller step-like increase in thin oxide and the QB

in thicker oxide, we have termed it interface trap-enhanced tunneling (ITET) due to its close correlation with interface trap generation. Subsequent stressing of the thin oxide results in much larger change in the gate voltage with magnitude similar to the occurrence of QB in thicker oxides.

Using the DCIV technique [6.30],[6.33], it is possible to monitor the evolution of bulk and interface traps at the channel region. Figure 6.3(a) shows the evolution of gate voltage V_g under constant current stressing for 45 Å oxides while Fig. 6.3(b) shows its associated DCIV measurement taken at the indicated interval as shown in Fig. 6.3(a). After onset of QB, it can be observed that the gate voltage decreases continuously with a certain period of temporary electrical recovery in gate voltage to higher level, although still lower than before QB. Since the number of bulk traps is proportional to the lateral shift in the peak I_B, it can be concluded that hole trapping occurs during the initial stage of stressing till QB. Moreover, since the peak amplitude of I_B (which is proportional to interface traps) is increasing, interface traps are also being generated at the same time. At quasi-breakdown (QB), it is observed that the DCIV spectra overlap with no continuous increase in both bulk and interface traps. Although post-QB oxides show different gate voltage under constant current stress, it can be observed that the peak I_B remains relatively constant (Refer to Fig. 6.3(b)). The result is similar to that obtained previously [6.3] and reaffirms that interface traps N_{IT} is not correlated to QB leakage current in 45 Å oxides.

For the thin oxide (20 Å), however, QB is observed much earlier with a smaller gate voltage fluctuation as commonly observed by other researcher groups [6.36]. After the first occurrence of QB, subsequent gate leakage for the thin oxide (20 Å) increases in steps, characteristic of further occurrences of QB spots [6.34]. Fig. 6.4 shows the associated direct-current current- voltage (DCIV) measurements of the (a) thick and (b) thin oxide samples. From Fig. 6.4(a), DCIV spectra for thick oxide (45 Å) shows continuous increase in the base current I_B with a left shift in the gate voltage at peak I_B, $V_{g,max.}$ Since the peak base current $I_{B,max}$ is directly related to interface traps while oxide trapped charge is proportional to $V_{g,max}$ [6.33], both interface traps and oxide trapped charges are increasing during Fowler-Nordheim

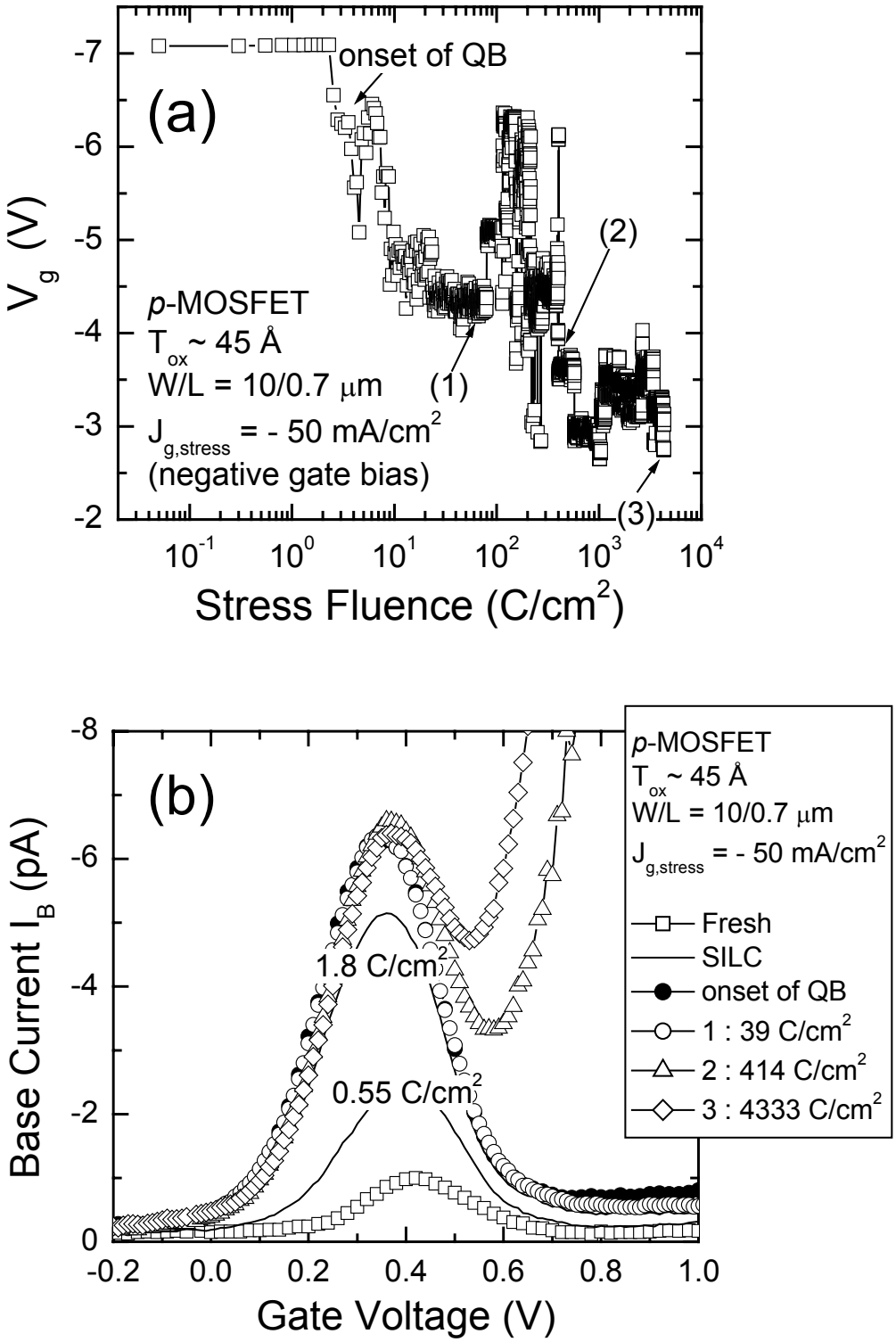


Fig. 6.3 (a) Evolution of gate voltage under CCS for thick (45 Å) p-MOSFET oxide after onset of QB. (b) Associated DCIV spectra measured at interval specified in (a). Base recombination current increases continuously under stressing till QB. At QB, DCIV spectra observed to overlap with no further increases. (Channel Area = 10 μm^2, p-MOSFET).

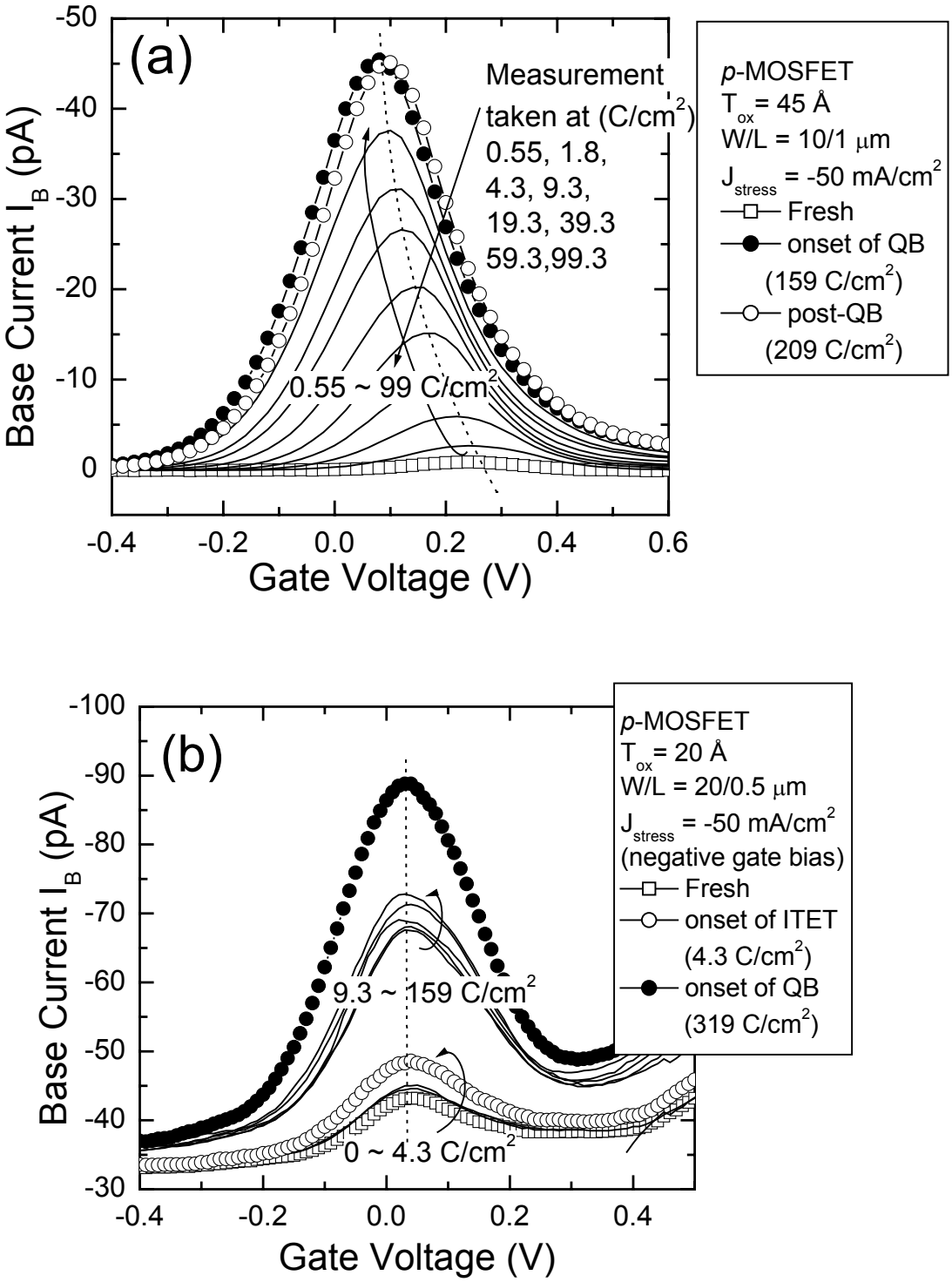


Fig. 6.4 Comparison of DCIV spectra for (a) thick (45 Å)and (b) thin oxide (20 Å) p-MOSFET under constant current stressing till QBs. (a) For thick oxide (45 Å), base recombination current increases continuously under stressing till QB. At QB, DCIV spectra observed to overlap with no further increases. (b) For thin oxide (20 Å), DCIV spectra increases even after QB with step-like increases in correlation with gate leakage current. (Locus of $I_{B,max}$ and $V_{g,max}$ is shown by the dotted lines). (Channel Area = 10 μm^2, p-MOSFET).

(F-N) constant current stress in thick oxide (45 Å). At QB, however, both interface and oxide bulk traps remains almost constant even after further continuous stressing. In thin oxide (20 Å), it can be observed from Fig. 6.4(b) that peak base current does not saturate after QB but increases in step-like fashion in exact correlation to gate leakage current. Since step-like increase in post-QB gate leakage current has been attributed to further generation of QB spots [6.34], it appears that the number of interface traps in thin oxide (20 Å) increases in tandem with further generation of QB spots. For thin oxide (20 Å), charge trapping is almost negligible as shown by the constant $V_{g,max}$ in Fig. 6.4(b). This is distinctly different from the case of thicker oxide, indicating that the QB mechanism in thick and thin oxides are distinctly different. A different mechanism which can be attributable to interface traps, which we named 'Interface Trap Enhanced Tunneling' (ITET), is responsible for gate leakage increases in thin oxides.

In thicker oxide (45 Å), it was previously reported in Chapter 4 and by other researchers [6.2], that the QB leakage current can be temporarily recovered by the application of a reverse bias anneal. This electrical recovery was observed together with the reduction in oxide trapped charge N_{OT} while interface traps remain constant. The results suggest that QB in thicker oxide is due to bulk traps rather than interface traps. However as oxide thickness is reduced to 20 Å, bulk charge trapping becomes insignificant. Moreover, the occurrence of QB is not always distinct and ITET appears to be governed by interfacial effects rather than bulk traps.

6.3.2 Direct Correlation between interfacial traps and gate leakage

It has been shown in the previous figure, Fig. 6.4 that ITET in thin oxide is governed by interfacial traps rather than bulk traps. Fig. 6.5(a) shows the direct correlation between gate leakage current and base recombination current $I_{B,max}$ (which directly reflects the interface trap density N_{IT}) during constant voltage stressing on 20 Å oxide using negative gate bias stress (gate injection). Compensated $I_{B,max}$ (= $I_{B,max}$ - $I_{g,,max}$) is used to account for gate leakage which may be significant for thin oxides. The result clearly shows that the increase of interface trap density under a negative gate bias stress occurs in a discrete manner and this is directly correlated to the gate

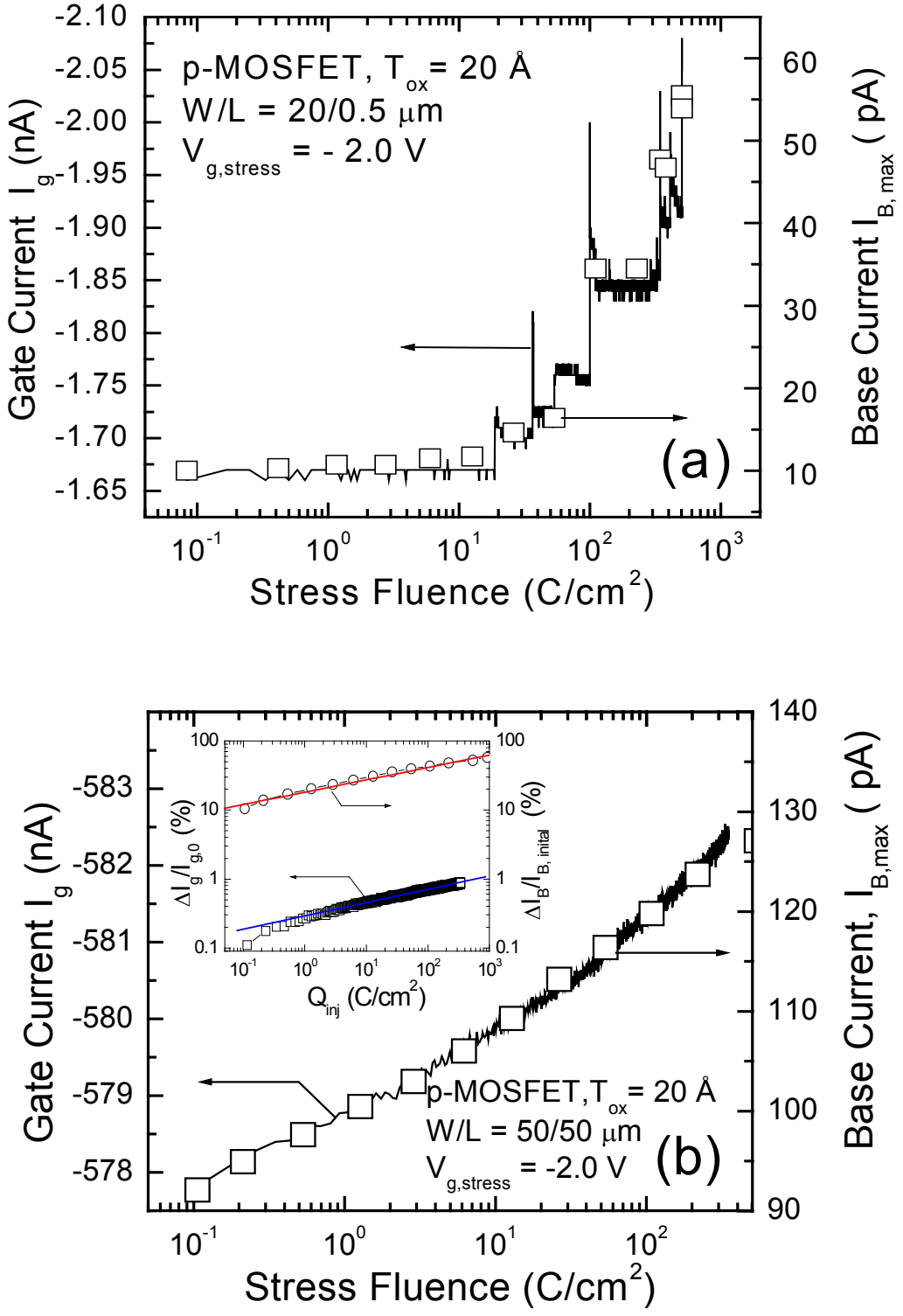


Fig. 6.5 Correlation of gate leakage current I_g and base recombination current I_B, which directly reflect interface trap density N_{IT} under constant voltage stressing for (a) small area samples (W/L = 20/0.5 μm) and (b) large area samples (W/L = 50/50 μm). Inset figure shows the percentage change in I_g and I_B. (T_{ox} = 20 Å, p-channel MOSFET).

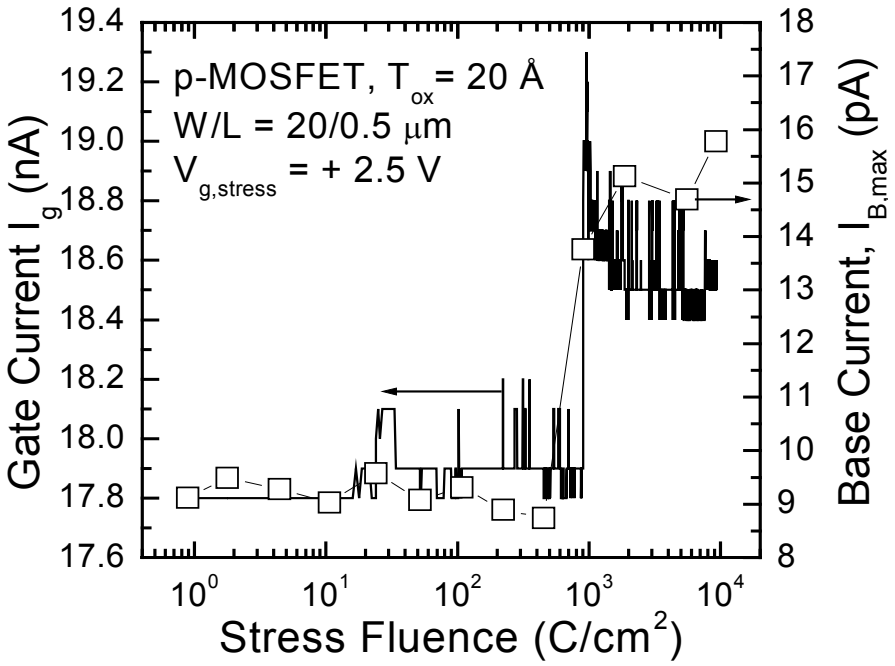


Fig. 6.6 Correlation of gate leakage current I_g and peak base recombination current $I_{B,max}$ under constant voltage stressing for substrate injection. Similar to negative gate bias (gate injection), gate leakage current observed to bear one-to-one correspondence to base recombination current which is directly correlated to interface traps. (T_{ox} = 20 Å, W/L = 20/0.5 μm, p-MOSFET).

leakage current when the gate oxide is stressed in the direct tunneling regime. Fig. 6.5(b) shows the evolution of gate leakage and peak base recombination current $I_{B,max}$ (inset figure shows the percentage change) under a constant voltage stress for large channel area samples (W/L = 50/50 μm). Unlike the small channel area (W/L = 20/0.5 μm) samples, the discrete increments are totally masked out in the larger area samples although a similar correlation can be observed in the inset figure as shown by the parallel lines between the percentage change in gate leakage and interface trap generation. The same slope for the percentage changes in gate leakage current and base recombination current (which reflects the level of interface traps) are obtained in the inset of Fig. 6.5(b), showing that both parameters are not simply increasing but are directly related.

Figure 6.6 shows similar correlation of interface trap to gate leakage current under positive gate bias stress. The results from Fig. 6.5 and Fig. 6.6 show that the occurrence of QB in thin oxides (20 Å) results in step-like increases in interface traps irrespective of the polarity of gate stressing. It also suggests that the generation of the

interface traps under a low constant voltage stress of less than 3 V (above 3 V, discrete increase in I_g is masked by the high gate direct tunneling current) is highly localized at certain spots similar to QB. The discrete increase in both gate leakage current and interface traps indicate that oxide degradation in thin 20 Å oxides is highly localized and directly correlated with interfacial traps formation. More importantly, since DCIV measures predominantly the interface traps only at the SiO_2-Si interface, the results suggest that the increase of the gate leakage current under an electrical stress whereby direct tunneling (DT) ($|V_g|$= 1.5 V to 3 V) dominates, is due mainly to interfacial degradation at the SiO_2-Si substrate interface. This is consistent with the fact that there are more strained bonds at the SiO_2-Si interface than the polySi-SiO_2 interface due to the oxide transition region created during oxide growth [6.37].

6.3.3 Distinction between ITT and ITET

The strong correlation between interface trap generation and gate leakage current suggests that the dominant mechanism of the gate leakage current degradation for the oxides in the direct tunneling regime is due to interface trap enhanced tunneling (ITET). The stressing and monitoring gate voltage used in this experiment is 1.5 V to 3.5 V and is much larger than the flat band voltage whereby interface trap tunneling (ITT) can be observed. Moreover, unlike interface-trap tunneling (ITT) which is due to tunneling current through interfacial traps and is observed at or near flatband condition, the ITET degradation is observed throughout the gate voltage range and is not confined to only flat band condition as shown in Fig. 6.7. Therefore, neither the band-to-interface trap tunneling model nor interface trap-to-interface trap tunneling model can be used to interpret the ITET current. In this case, it is observed that similar inference has been drawn by Meinertzhagen *et al.* [6.15] who concluded that low voltage SILC (LV SILC) cannot account for all the degradation observed at other gate bias away from the flatband conditions.

Figure 6.8 shows the carrier separation studies for hole and electron current of p-MOSFET gate oxide stressed under a negative gate bias. Under channel inversion, electron and hole currents can be separately monitored through well and source

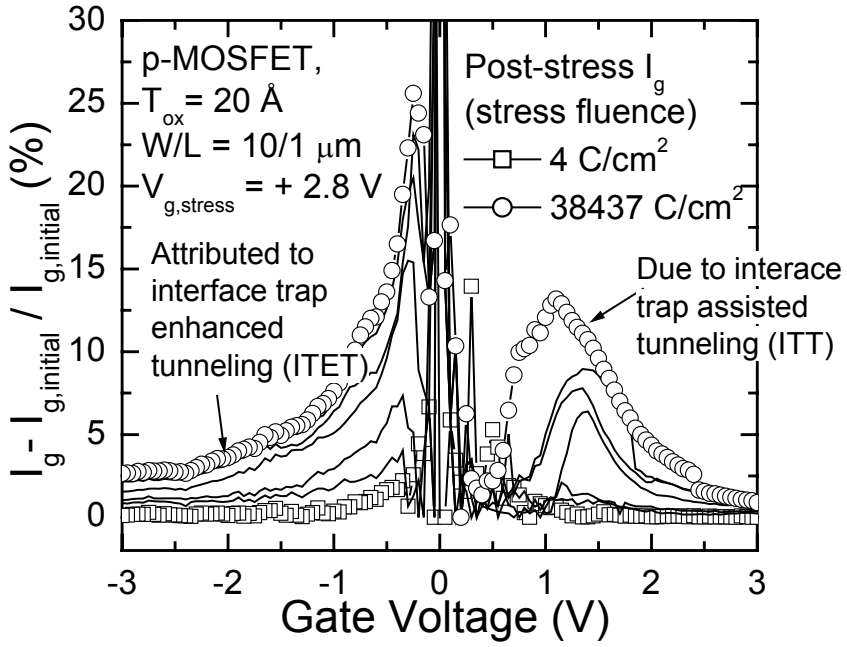


Fig. 6.7 Percentage change of gate leakage current (I_g–$I_{g,initial}/I_{g,initial}$) for different gate voltage under constant voltage stressing. The spike for ΔI_g at gate voltage near to zero volts is due to background noise. Unlike ITT, ITET occurs throughout the entire voltage measurement range. (T_{ox} = 20 Å, Area = 10 μm^2, p-channel MOSFET).

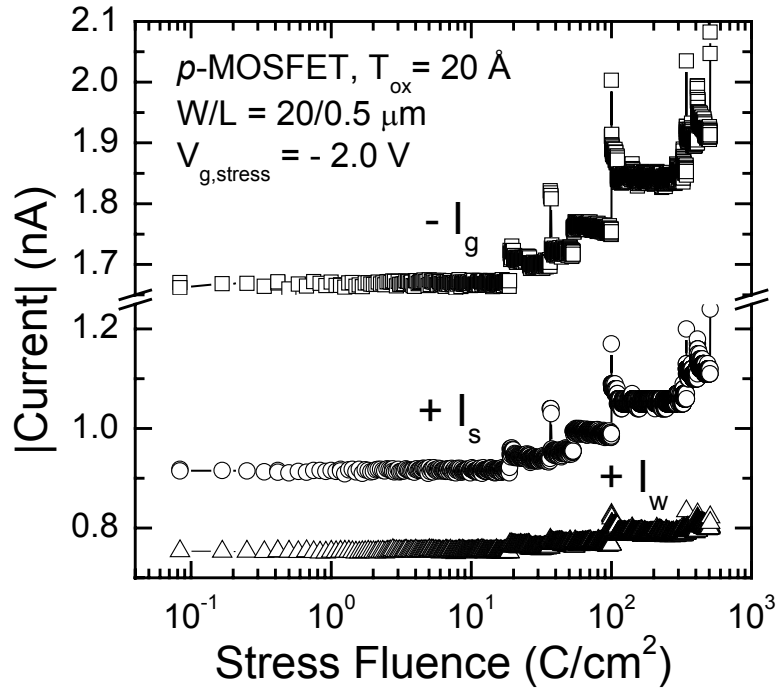


Fig. 6.8 Carrier separation showing holes (shown by source current I_s) and electrons current (shown by well current I_w) for small channel area p-MOSFET stressed under negative constant gate voltage. (T_{ox} = 20 Å, Area = 10 μm^2, p-channel MOSFET).

current respectively. It can be observed that the discrete increase in the gate leakage current is due to a similar discrete increment in both the electron and hole currents. The result shows that the discrete increase in gate leakage is due to an enhancement in both the holes and electrons direct tunneling currents. Based on the result presented, local interface trap formations appear to lead to an enhancement of the hole and electron direct tunneling which can be attributed to either equivalent oxide thinning or energy gap lowering mechanism at the local spot due to very high density of interface traps.

6.3.4 Area dependency of ITET

Using different channel area, the area dependency of ITET was also studied and is shown in Fig. 6.9(a) and (b). It can be observed that when the gate current density is used as device failure criteria, smaller sample area shows higher increases in current density although the actual current increases as shown in Fig. 6.9(b) is actually smaller. When a larger sample area is used as shown in Fig. 6.9(a) (Refer to 2500 μm^2 sample), the discrete increase in gate leakage current is not observable, but instead a gradual increase is observed. The result shows that the localized degradation leading to ITET does not have a uniform area density but is also not entirely independent of area. In our studies, we have deliberately selected a wide range of channel area ranging from 10 μm^2 to 2500 μm^2. As the localized degradation spot will be relatively small compared to the channel area, it is important to include smaller area to accentuate the comparison. As observed in Fig. 6.9(b), an area dependency can be observed which accounts for the increasing leakage current for larger channel area. The results refute the claim that gate leakage current due to progressive wearout is independent of area.

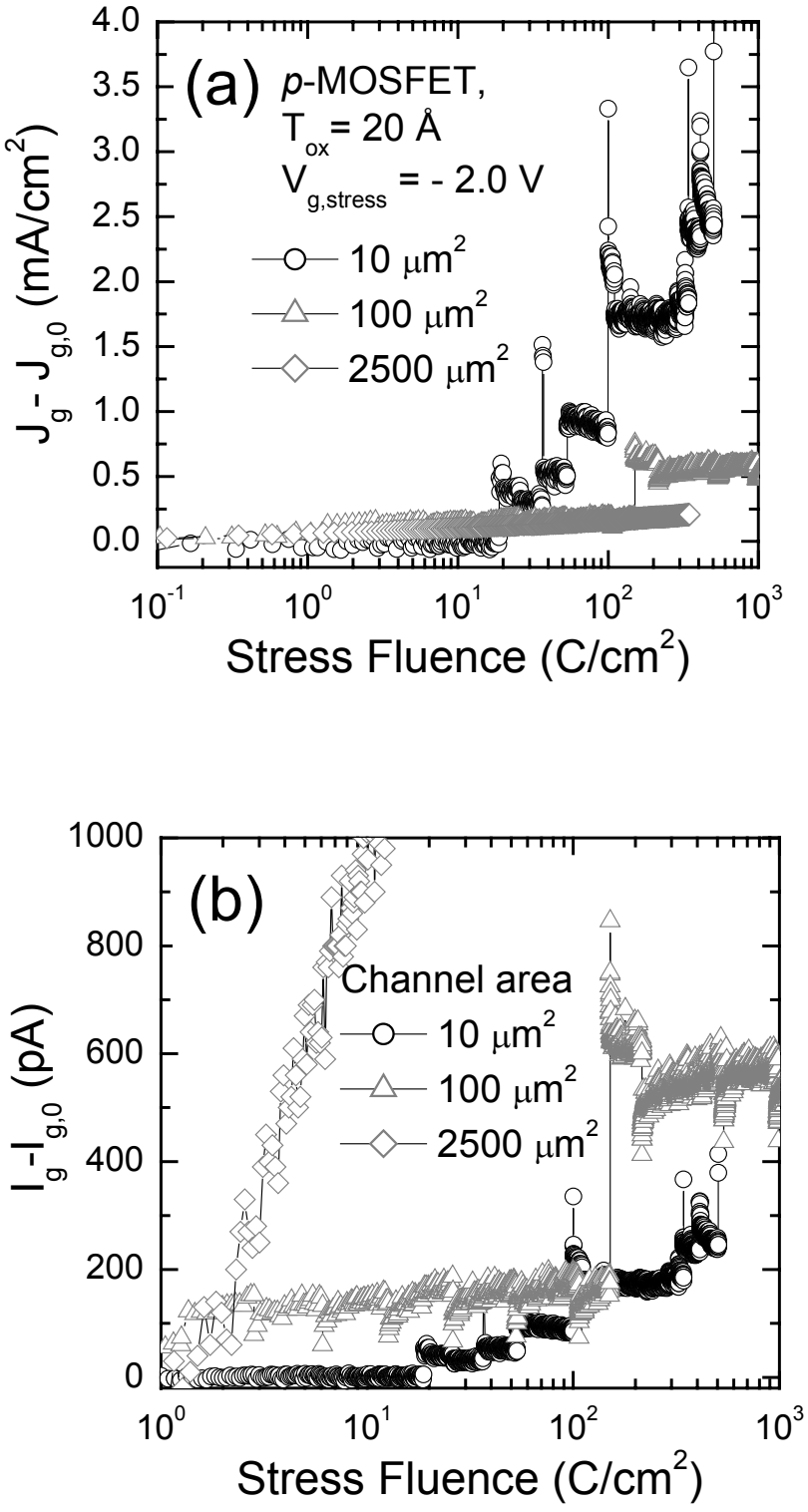


Fig. 6.9 (a) Discrete gate current density increase for different channel areas (b) normalized gate current increase under constant voltage stress for different channel area ranging from 10 to 2500 μm^2. (T_{ox} = 20 Å, p-MOSFET).

6.4 Ultra-thin (<14 Å) oxide Degradation Mechanism

As the thickness of gate oxides is further scaled, it was observed that QB, or any form of discrete gate leakage increases becomes non-existent. Instead a progressive wearout is observed. In most cases, for oxides thinner than 20 Å, QB is defined as the onset of voltage/current fluctuation in the gate bias/current under CCS/CVS and cannot be distinctly observed except by resolving its voltage/current noise parameter [6.16]. In the next few sections, the oxide degradation for 13 Å oxides is described. Figure 6.10 shows the current-voltage (I-V) characteristics of p-MOSFET with source/drain and n-well grounded. In the positive gate voltage regime, for gate voltage V_g < 1 V, hole current from p+ gate to p+ source/drain dominates. This is due to the similar doping type of both p+ gate and p+ source/drain resulting in the alignment of the p+ gate Fermi level with the p+ source/drain valence band at low negative gate voltage. For V_g > 1 V, the Fermi level of the n-well surpasses the conduction band at the PolySi gate electrode allowing electrons to tunnel to the p+gate anode. As a result for V_g > 1 V, electron injection from n-well substrate dominates the gate leakage current. In contrast, for negative gate bias regime, hole

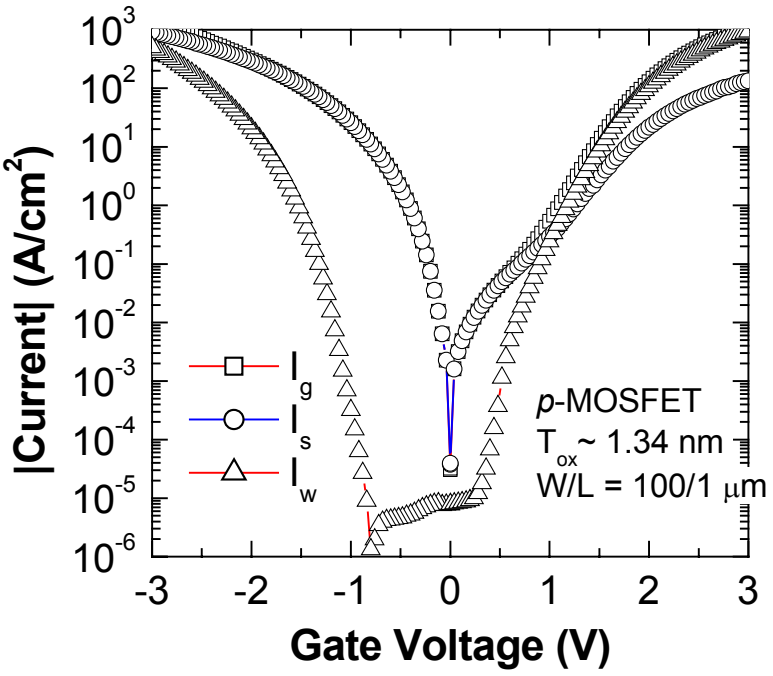


Fig. 6.10 Carrier separation characteristics for 13 Å gate oxide in both depletion and accumulation. Source, n-well and gate current indicated by I_s, I_w and I_g respectively. Drain electrode is not connected. (Channel Area = 100 μm^2, p-MOSFET).

current from source/drain dominates over valence electron injection from the gate, for low to moderate negative gate voltage regime (0 V to -2 V). At high negative gate bias, valence electron becomes significant although it is still lower than the hole current from p-source. In the case of negative gate stressing, both electron and hole injections are significant at high negative gate bias. In contrast, at high positive gate bias, electron injection is much higher than that of holes, resulting in primarily electron injection from n-well.

Figure 6.11(a) shows the gate current under a negative constant voltage stress ($V_{g,stress}$ = - 3.0 V) for ultra-thin (13.5 Å) oxide. After about 200 s of constant voltage stressing, it can be observed that gate current increases monotonically with logarithmic of stressing time. Under negative gate voltage with channel depletion in n-well (p-MOSFETs), valence electrons injected from gate will be channeled to the n-well while holes from the source/drain are injected into the gate. Fig. 6.11(a) also shows the corresponding source and n-well currents. For all current-voltage (I-V) measurements, the drain is not connected. From Fig. 6.11(a), it can be observed that the increased gate leakage current is due primarily to hole current injection from source to gate while the increase in valence electron injection from p+ gate is less significant. This increased leakage current is significantly larger than the initial stress-induced leakage current and should be considered as a different phase. We ascribed this progressive increase in gate leakage current or second phase of increased leakage current to multiple quasi-breakdown spots in the thin gate dielectric. Unlike complete breakdown, it is noted that the current-voltage leakage characteristic of the degraded oxide after the onset of progressive breakdown (Progressive BDs) is still significantly smaller than that after complete breakdown.

Fig. 6.11(b) shows the evolution of the current-voltage (I-V) characteristics of the gate oxide under constant voltage stress V_g = -3 V with electron gate injection. The I-V curves are taken at equal logarithmic time interval. In the negative gate regime, where –0.8V < Vg < 0V, gate leakage current shows no increase which corresponds to the absence of valence electron injection from gate due to alignment of the forbidden bandgap in the gate and substrate. In all other gate voltage regime, gate

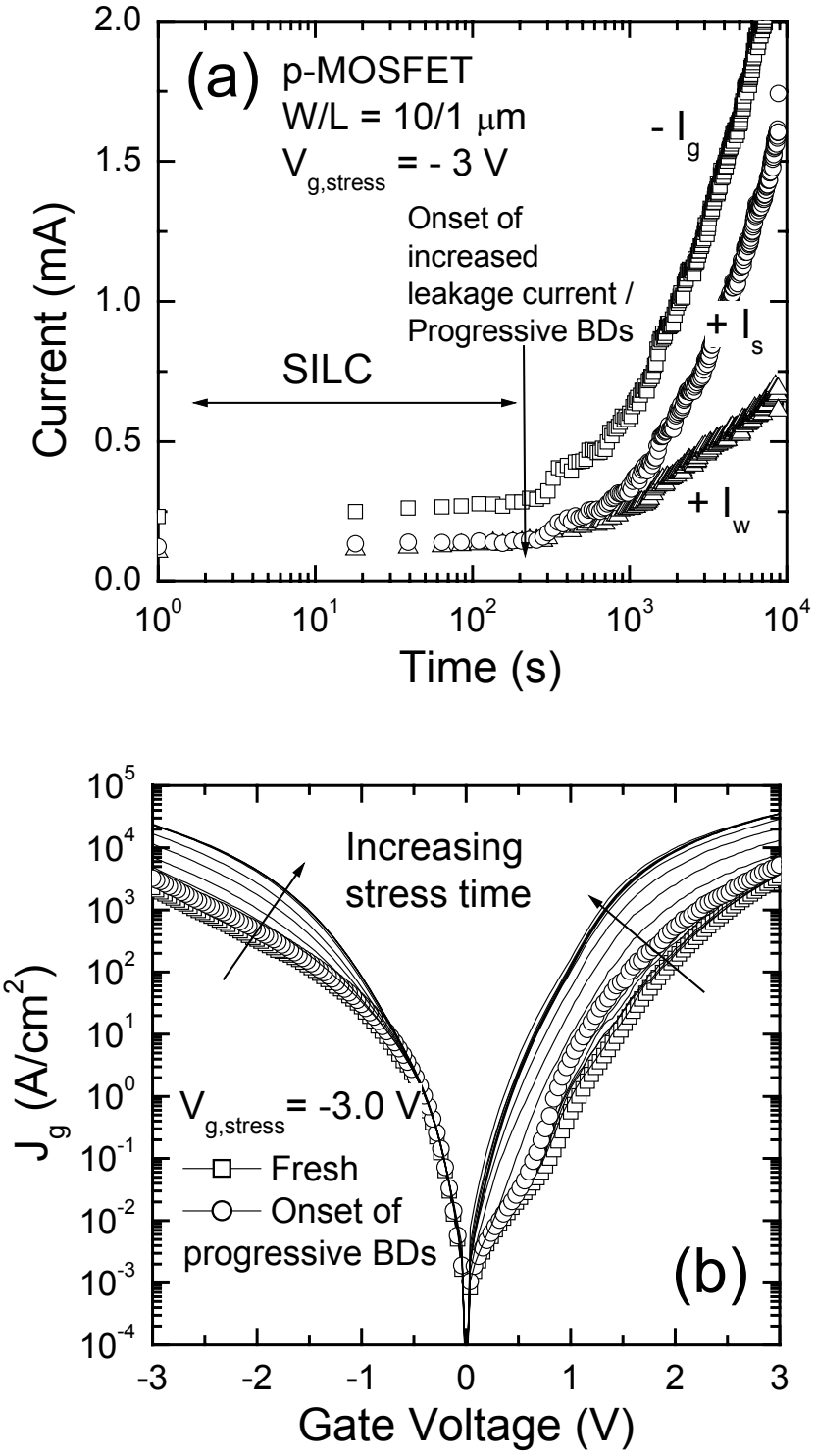


Fig. 6.11 (a) Evolution of current-voltage characteristics at fresh and after constant voltage stressing (V_{stress} = -3.0 V). (b) Associated I-V characteristics at onset of progressive BD and post PBDs. (T_{ox} = 13.4 Å, Channel Area = 10 μm^2, p-MOSFET).

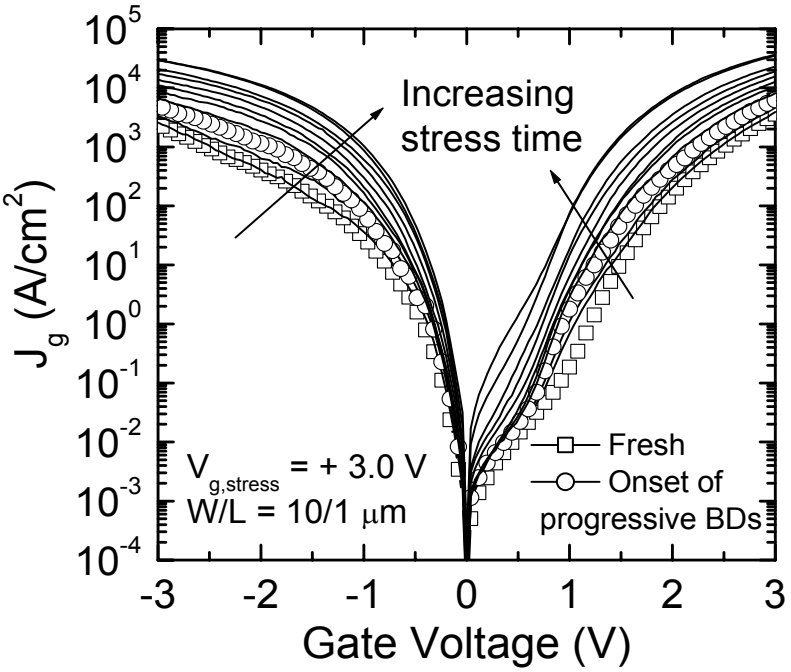


Fig. 6.12 Evolution of current-voltage characteristics at fresh and after constant voltage stressing (V_{stress} = + 3.0 V). (T_{ox} = 13.4 Å, Channel Area = 10 μm^2, p-MOSFET).

leakage current shows significant increase after onset of the increased leakage current or progressive BDs phase. For positive gate voltage stressing, electron current from n-well increases significantly (not shown here) after onset of the increased leakage current phase similar to that observed for negative gate stressing.

Figure 6.12 shows the associated current-voltage characteristics of the p-MOSFET under positive gate bias stress. The main difference for post-positive gate stressed I-V characteristics from that of negative gate stressed as shown in Fig. 6.10(b) is the existence of an anomalous current increase for 0V < V_g < 1 V. This corresponds to the low-voltage enhanced interface trap-assisted tunneling near to flatband conditions as described by A. Ghetti *et al*. [6.13]. However, similar to the case for negative gate bias stress, no discrete increase in gate leakage current is observed under a positive gate bias stress. Instead, gate leakage increases progressively even in small area channel devices.

Figures 6.13(a) and (b) show the evolution of normalized gate leakage current under electrical stressing with different constant gate voltages. Normalized gate leakage current is defined as $\Delta I_g/I_{g,0}$ where $\Delta I_g = I_g - I_{g,0}$ and $I_{g,0}$ and I_g are the gate

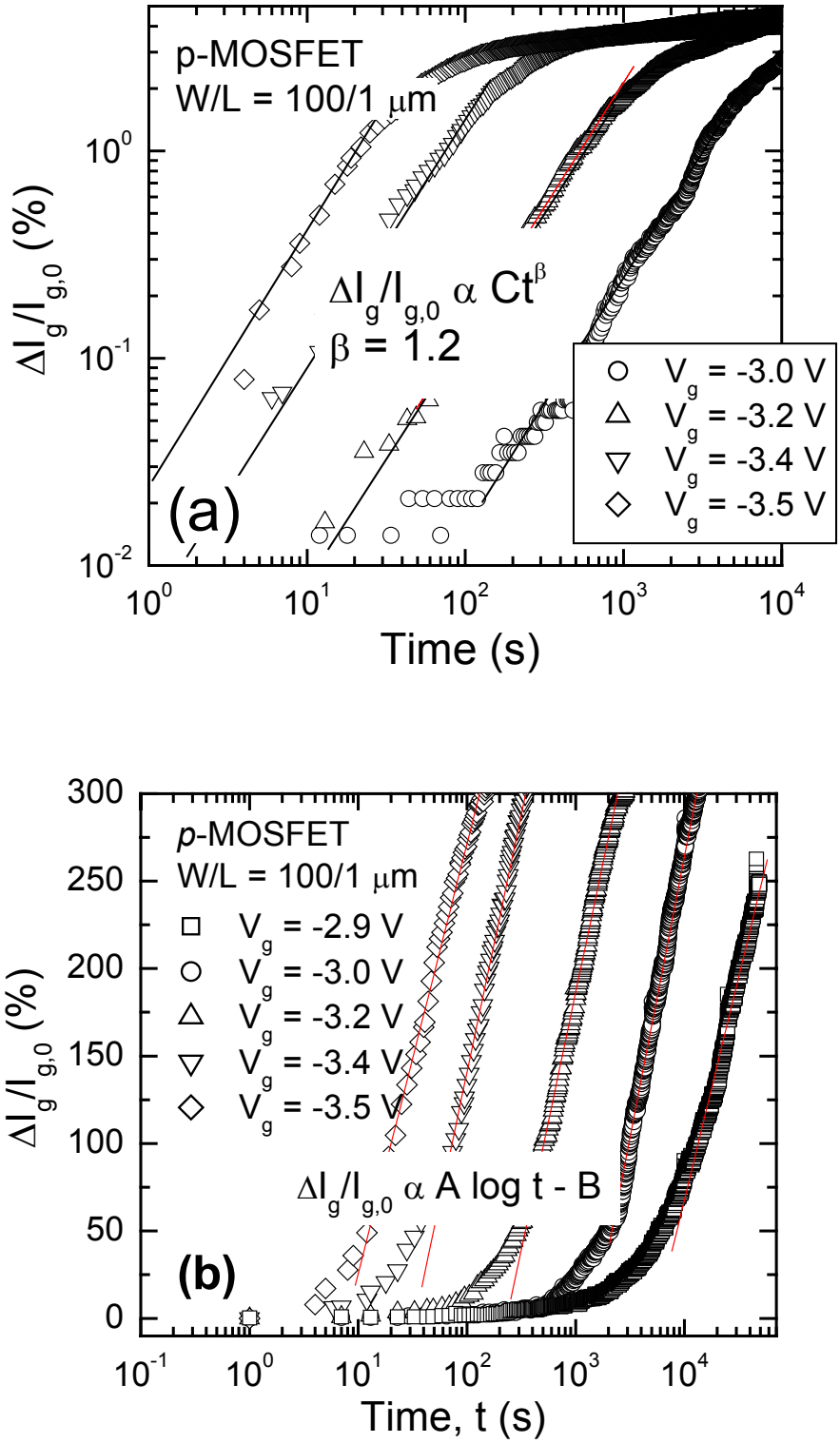


Fig. 6.13 (a) Percentage change in gate leakage current under different stressing gate voltage in the initial stage of PBDs. (b) In the subsequent PBDs stages, leakage current proportional to logarithmic of stressing time. It can be observed that gate leakage current follow a power relation with stressing time in the initial stage (a) and a linear logarithmic time dependence as shown inset in the subsequent stage (b).

current before and after stress respectively. It is observed that the onset of PBDs is characterized by two stages in degradation characteristics. In the first stage, gate leakage current increase follows a power relation with stressing time with power factor of β = 1.2 as shown in Fig. 6.13 (a). As stressing continues, gate current increases follow a logarithmic time increase as shown in Fig. 6.13 (b). An almost parallel shift is observed in normalized gate leakage currents for different gate stress voltages at the two different stages of PBDs but with different relations to stressing time as shown in Fig. 6.13 (a) and (b). In the initial stage, gate leakage current increases can be modeled by uncorrelated multiple occurrences of breakdown spots. Alam *et al*. have shown that by assuming an uncorrelated breakdown (BD) event, gate leakage current under several BD spots, follows a power relation with stressing time [6.24], which is exactly what is observed in Fig. 6.13(a). This is also further verified in 6.5.2 whereby the gate leakage current in the initial stage can be ascribed to multiples occurrence of uncorrelated BD spots. Eventual stressing will continue to a stage whereby all the undamaged region of the gate dielectrics are completely damaged. At this point, no further generation of new BD spots are possible and gate leakage current should saturate. This is shown by the second stage as shown in Fig. 6.13(b), whereby gate leakage current increases are limited and increases at a significantly lower rate with logarithmic time dependence. The nominal gate leakage current increase in the second stage suggests that even at the degraded spots, further degradation is possible although this effect is much less significant on gate leakage current than that of the generation of new BD spots.

Figure 6.14 shows the plot of the slope parameters *A* and y-intercept parameter *B* obtained from Fig. 6.13(b) for the localized gate leakage current increase during the 2nd stage of PBDs. It can be seen that normalized gate leakage current can be described by the following equations where A = 2.7 and $B = C$ -12.85 $|V_g|$:

$$\frac{\Delta I_g}{I_{g,0}} \approx A \log t - B \tag{6.1}$$

where V_g is the applied gate voltage (in volt) and C is a constant. The voltage acceleration obtained in this case is 12.85 and this is quite close in magnitude to the value of 12.5, obtained by Monsieur *et al.* [6.21].

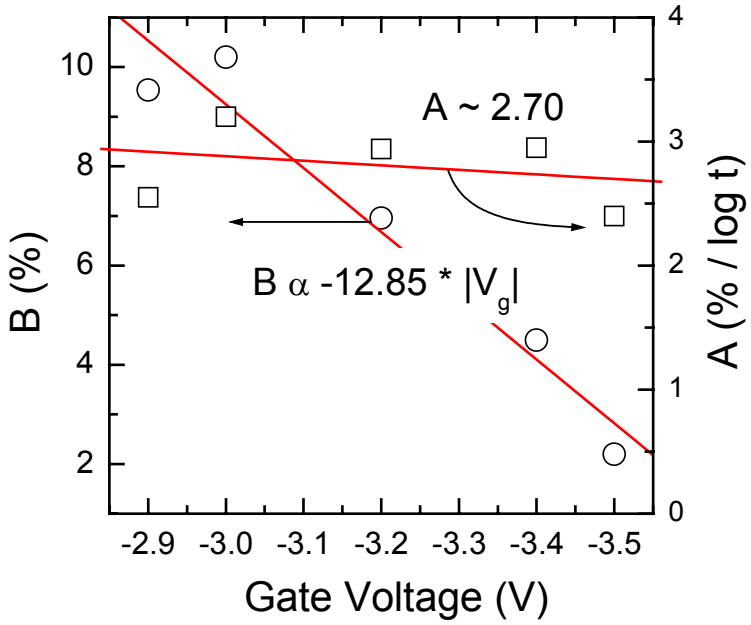


Fig. 6.14 Graph showing gate leakage current time dependence parameters A and voltage dependence parameter, B as defined in (1) for 2nd stage of PBDs versus stressing gate voltage. (Channel Area = 100 μm^2, p-MOSFET).

6.4.1 Area Dependency of Progressive Breakdown in Ultra-thin Oxides

In thick oxides, stress induced leakage current (SILC) results in uniform degradation, independent of channel area, while QB or ITET shows significant area-dependency due to its discrete localized breakdown mechanism. In retrospect, area dependency studies will help us ascertain the nature of progressive breakdown observed in ultra-thin oxides. Fig. 6.15 (a), (b) and (c) show the current density J_g , normalized gate current $\Delta I_g = I_g - I_{g,0}$ and percentage change $\Delta I_g / I_{g,0}$ for different sample area when a constant voltage of –3 V is applied to the gate. (Negative sign for J_g indicates that valence electron is flowing from the gate electrode to the substrate). From Fig. 6.15(a), it can be observed that the increased gate leakage current phase does not scale uniformly with device area. Instead, there is a significantly higher leakage current density for smaller device area when compared to larger sample area. In contrast, gate leakage current defined by the gate current change $I_g – I_{g,0}$ which shows the magnitude of gate leakage current increases for different area shows an increasing leakage current for bigger sample area with eventual saturation for 100 μm^2 or larger area as shown in Fig. 6.15(b). This discrepancy between the current

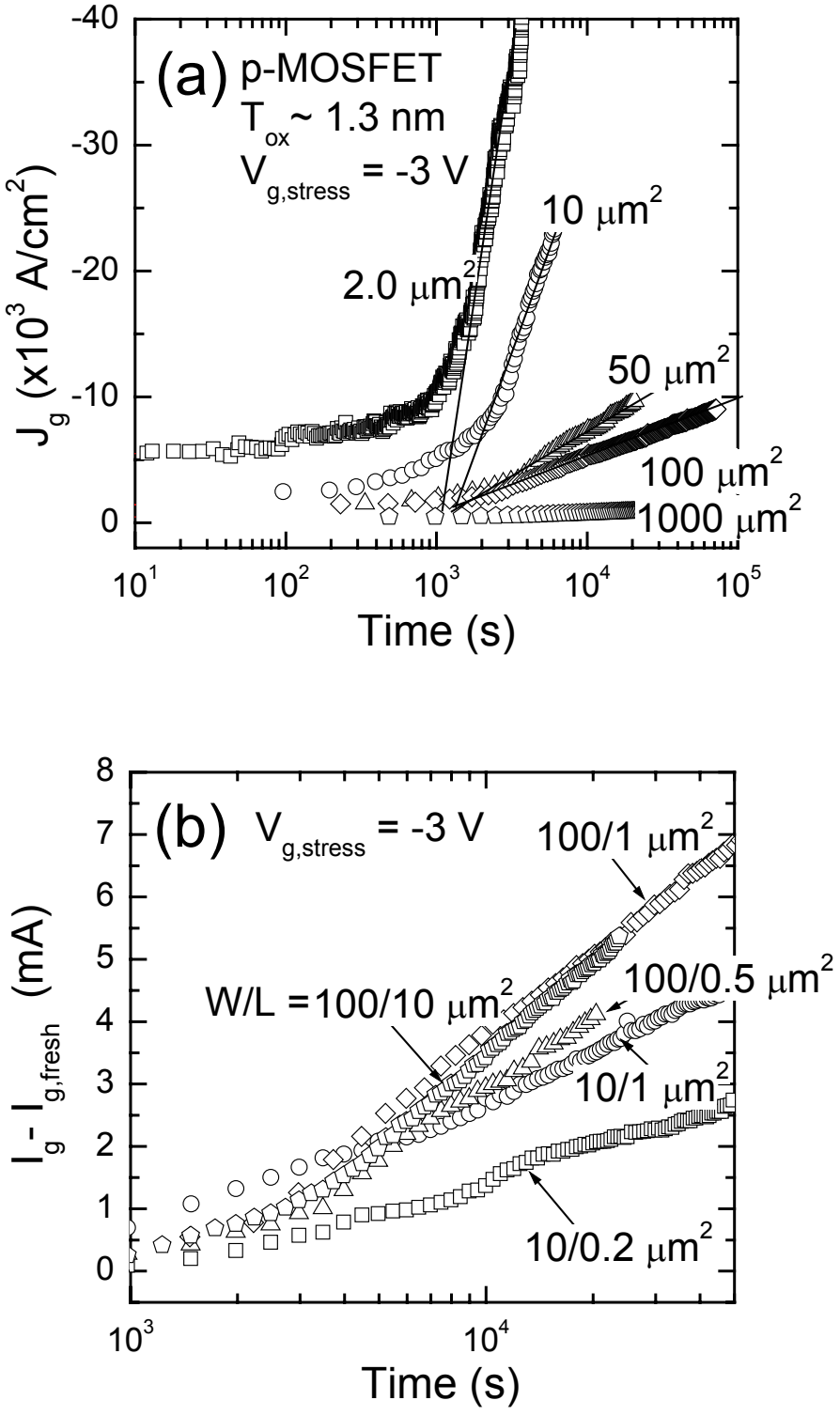


Fig. 6.15 (a) Gate leakage current density increases under constant voltage stress of V_g = -3. 0 V for different sample areas (b) absolute increase in gate leakage current for different sample area stressed under CVS..

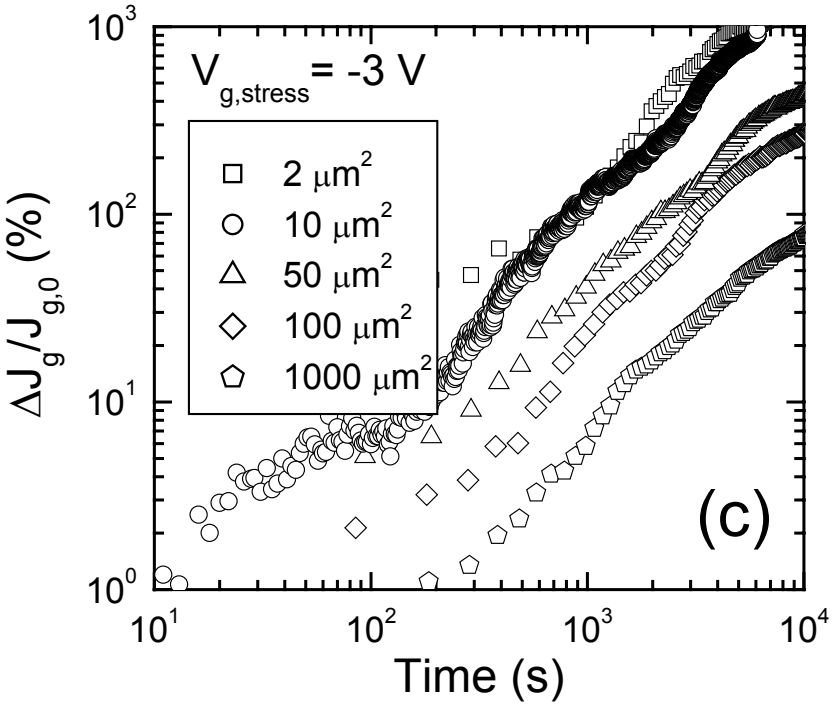


Fig. 6.15 (c) Percentage increase in gate leakage current after onset of 2nd stage PBDs for different sample areas.

density and absolute current increase can be reconciled if the degradation mechanism is highly localized with certain parametric area, similar to that of quasi-breakdown or ITET observed in thin oxides. The area dependency for the progressive leakage current shows varying gradients for different areas when plotted with logarithmic time as shown in Fig. 6.15 (a). It can be noted that parameter *A* defined in Fig. 6.14 and Eq. (6.1) has a direct area dependency. This lack of temporal distribution for different device area in Fig. 6.15 (a) suggests that each progressive BDs spot is exceedingly small such that its first occurrence is probably masked out by the background high direct tunneling current. Fig. 6.15(c) shows that in the initial PBDs stage, area scaling can be represented by a parallel shift in the leakage current in the time domain. This verifies that the onset of leakage current is actually different in the different area and obeys the Weibull scaling rules as will be shown later.

Figure 6.16 shows the defect generation rate or degradation rate of oxide defined by the increase in gate current density over injected fluency and the time-dependent parameter *A,* defined in Eq. (6.1) which has an area dependency as discussed previously. From Fig. 6.16, it can be observed that degradation rate P_g is

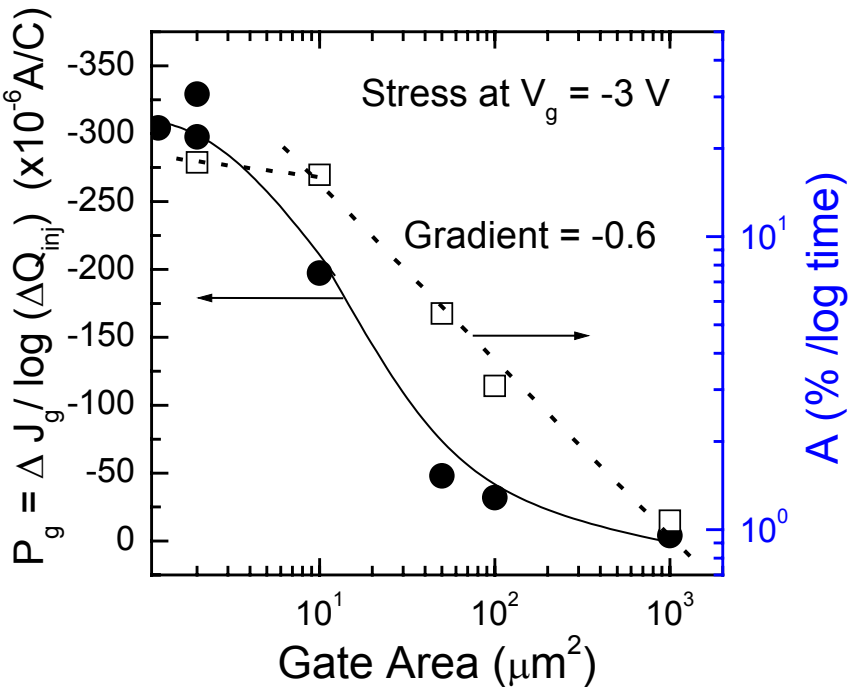


Fig. 6.16 Degradation rate or defect generation rate defined by $J_g = P_g * Q_{inj}$ where P_g is the defect generation / gate degradation rate as shown for second stage of PBDs. It can be observed that gate degradation increase as sample area decreases with eventual saturation at very small area.

higher in smaller area compared to larger area devices. This is due to the discrete nature and localization of each progressive BDs which tends to influence smaller area more severely. Unlike conventional Weibull's statistics which predict a longer time-to-breakdown for smaller gate area, onset of progressive BDs result in a higher gate leakage current density for smaller area devices. At the same time, it can be noted that the parameter *A* for the 2nd stage of PBDs can be expressed as a function of gate area with power proportionality of -0.6 of the gate area with eventual saturation for device with gate area smaller than 10 μm^2. The negative slope value of -0.6, means that smaller areas will have a proportionally steeper increase in gate leakage current, as compared to bigger samples. The implication of this is shown in Fig. 6.17. In this case, 2 criteria are considered for lifetime projection. In the first criterion, time-to-100% increase, $t_{100\%\ Ig}$, in leakage current is considered as lifetime of the device. In contrast, the time-to-complete breakdown, t_{BD}, shows the time to failure of the oxide due to complete breakdown in the oxide. From Fig. 6.17, it can be observed that the conventional t_{BD} is far too optimistic when compared to an increased leakage current

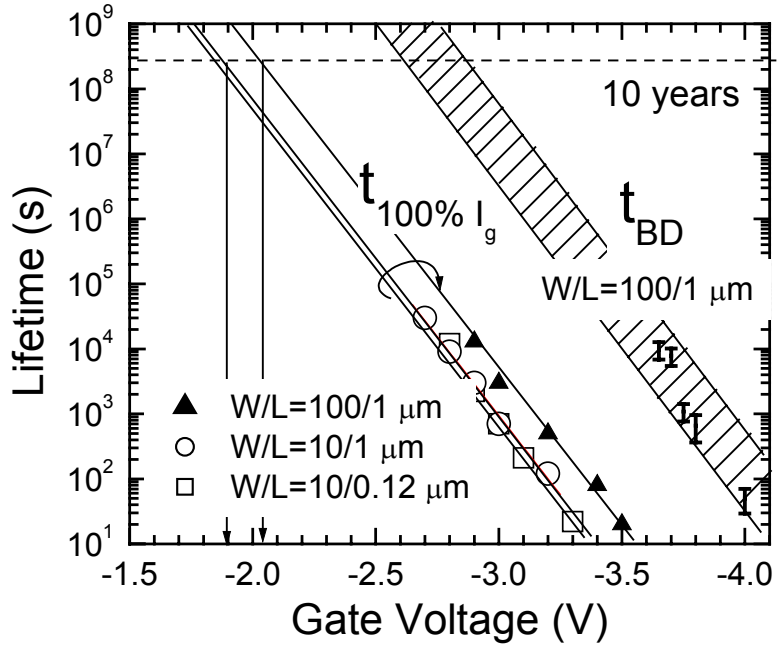


Fig. 6.17 Lifetime projection versus stressing gate voltage using 100% increase in gate leakage current as failure criterion. Lifetime for 100% increase in gate leakage $t_{100\% Ig}$ is much shorter than the conventional time-to-complete breakdown t_{BD}. (T_{ox} = 13.4 Å, p-MOSFET).

criterion. Moreover, the localization nature of the progressive breakdown resulting in a lower defect generation as shown in Fig. 6.17 meant that bigger area is actually better in term of gate leakage degradation. As shown in Fig. 6.17, with the new criterion, unlike the conventional time-to-QB or complete BD, $t_{100\%Ig}$ does not have a constant area scaling factor in the Weibull plot, and instead, actually favors large samples. This is a consequence of the impact of PBD on different sized samples which shows greater impact in smaller sized samples. Using this new criterion, the operating voltage with projected 10-year lifetime V_{10Y} is about 2.03 V for 100 μm^2 gate area and 1.9 V for 10 μm^2 gate area.

6.5 Modeling of Gate Leakage Current in Ultra-thin Oxides

From the experimental data presented in this study, an empirical relationship is obtained relating gate stress voltage, time and area dependency to the progressive leakage current component in the following sections:

6.5.1 Empirical Experimental Fitting

From Fig. 6.13, it was observed that at the initial stage of PBDs, gate leakage current increases follows a power relation with stressing time as shown below:

$$\frac{\Delta I_g}{I_{g,0}} \approx Ct^{\beta} \tag{6.2}$$

where C and β are constants. β = 1.2

At the subsequent stage of PBDs, gate leakage current increases logarithmically with stressing time :

$$\frac{\Delta I_g}{I_{g,0}} \approx (A)\log t + 12.85\left|V_g\right| - C$$
$$A = A_0\left[\frac{A_T}{A_{T,0}}\right]^{-0.6} \quad for Area, A_T > 10\mu m^2 \tag{6.3}$$

where A_T is the gate total area, $A_0 = 2.7$ and $A_{T,0} = 100\ \mu m^2$

6.5.2 Modeling of Multiples Breakdown Spots

To explain the area dependence observed in the localized gate leakage current, we have adopted a model consisting of multiple degraded spots within the ultra thin gate oxide. This localized phenomenon of the gate leakage current is not a unique characteristic but is also observed in breakdown and quasi-breakdown. Oxide breakdown studies, reported by other researchers, have shown that the dielectric breakdown is a localized event [6.1]-[6.5]. This is supported by emission microscopy images [6.1] capturing the dielectric breakdown event. In our modeling, we assume the followings:

1. Each degradation spot has similar area and current-voltage (I-V) characteristic and the resulting gate leakage current is a superposition of the fresh gate leakage current and that of multiple degraded spots.
2. The local degraded spot does not propagate with continual stress and that its I-V characteristic remains unchanged once the damaged state is reached.

3. That the occurrence of subsequent degraded spots is still random and not a lateral enhancement of the present degraded spots. In this respect, M. A. Alam *et al.* have shown that multiple soft breakdown is statistically independent [6.25]. This ensures that Poisson formalism is still applicable for each subsequent spots.

For a given gate area, A_T, the maximum possible number of degraded spot is given by $N = A_T/A_l$ where A_l is the area of each individual degraded spot. Similar to Lee and Cho *et al*.'s local physical damaged region model [6.1] in quasi-breakdown, the gate current density over the fresh and local degraded spot can be expressed as:

$$\begin{aligned} J_g &= J_{g,0}\left(\frac{A_T - n(t)A_l}{A_T}\right) + J_{PDR}\left(\frac{n(t)A_l}{A_T}\right) \\ &= J_{g,0}\left(1 - n(t)R_A\right) + J_{PDR}n(t)R_A \\ &= J_{g,0} + n(t)R_A\left(J_{PDR} - J_{g,0}\right) \end{aligned} \tag{6.4}$$

where J_g is the post-stress gate leakage current density, $J_{g,0}$ is the fresh gate leakage current density and J_{PDR} is the increased current density through the degraded oxide. $n(t)$ is the number of degraded spots and $R_A = A_l/A_T$ is the ratio of the area of each degraded spot versus the initial total sample's area. Applying the assumptions, R_A and R_J are then invariant with time as shown below, where R_J is the ratio of the current density of the degraded spot J_{PDR} over the fresh current density $J_{g,0}$. The normalized gate leakage current can then be simplified from Eq. (5.4) as shown below:

$$\begin{aligned} \frac{\Delta J_g}{J_{g,0}} &= \frac{J_g - J_{g,0}}{J_{g,0}} \\ &= n(t)R_A\left(R_J - 1\right) \end{aligned} \tag{6.5}$$

From Eq. (6.5), we can determine the approximation of $n(t)$ for both small and large area samples. For large area samples where $R_A \ll 1$, $n(t)$ can be considered as a continuous variable, $n(t) = F(t)/R_A$ where $F(t)$ is the cumulative probability frequency ($0 \le F(t) \le 1$) of the localized spots. This means that Eq. (6.5) can be rewritten as:

$$\frac{\Delta J_g}{J_{g,0}} = F(t)\left(R_J - 1\right) \tag{6.6}$$

From Eq. (6.6) and invoking Weibull's area scaling rule, we observe that the relation between the cumulative probability function for the spots F(t), and the entire sample, $F_T(t)$ is simply given by

$$1 - F(t) = (1 - F_T(t))^N$$
$$F(t) = 1 - (1 - F_T(t))^N \tag{5.7}$$

We note that Eq. (6.7) is actually a simple form of the area scaling for extreme value cumulative frequency distribution function and from Eq. (6.7), it can be seen that the total possible degraded spots N will affect the leakage current as ΔJ_g is proportional to F(t). Since N is proportional to the total gate area A_T, and assuming A_l remains the same, the leakage current percentage increase which is proportional to F(t), will be smaller for larger area sample as compared to smaller area sample. At the same time, since the Weibull slope β is not affected by the area scaling effect, the sample Weibull slope β can be derived directly from the spot's Weibull distribution.

Figure 6.18 shows evolution of normalized gate leakage current for p-MOSFET when stressed under gate voltage. Using Eq. (6.6), it can be seen that as time$\rightarrow \infty$, $F(t) = 1$ and R_J can be obtained. By using different R_j values, we are able to derive the Weibull plot given a certain gate leakage current graph. From Fig. 6.18, the

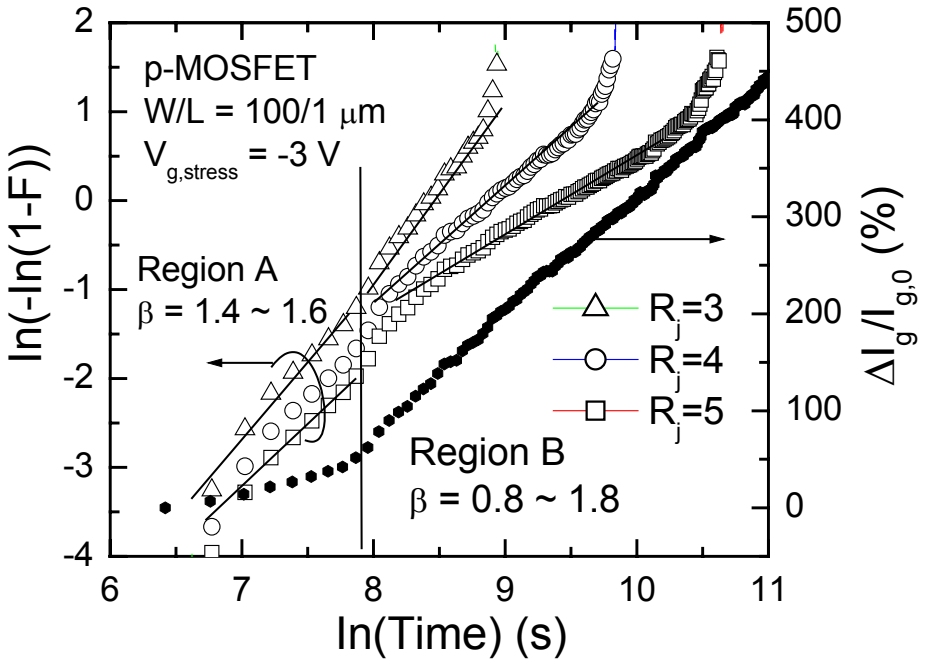


Fig. 6.18 Evolution of normalized gate leakage current $(I_g - I_{g,0})/I_{g,0}$ on a 13.4 Å gate oxide when stressed under constant voltage stress (as shown by the solid symbol). Using (5.6), the cumulative frequency of localized spots occurrence F(t) can be determined and it's associated Weibull plot (with various values of R_j) versus logarithmic of stressing time is as shown (shown by the various open symbols). It can be observed that the derived Weibull plots can be separated into 2 regions, A & B. The derived Weibull shows a good linearity when $R_j = 3$ for both region A and B, deviating only at the extreme short and long stressing time due to censoring effects.

derived Weibull plot using different values of R_j ranging from 3 to 5 is simulated. It can be observed that the derived Weibull plot can be divided into 2 regions, region A & B. When $R_j = 3$, very good linearity is obtained for the Weibull plot in both region A and B. The Weibull plot deviates from a straight line at the extreme short and long stressing time due to censoring effects, which has been thoroughly explored by Rowland *et al*. [6.38]. In this case, censoring effect is unavoidable at the early stage of the localized leakage current due to the initial presence of direct tunneling current, which masks out the initial localized leakage current occurrence. As the stressing is not till complete saturation and occurrence of all local spots, long time censoring is also unavoidable. The good linearity in this plot implies that the localized degraded spots can also be described by a Weibull's distribution. Since this simple analytical methodology measures the statistical distribution of degraded spots over the whole

device area, it also opens the new possibility of oxide degradation studies using limited number of samples.

6.6 Summary

In conclusion, gate leakage current in ultra-thin oxide (20 Å) is characterized by discrete step-like increases rather than gradual increases during constant voltage stressing, which can be attributed to occurrences of multiple QB. Correspondingly, the discrete step-like increases of the interface traps are also measured by the DCIV method and a very good correlation was observed between gate leakage current and interface trap density irrespective of stressing polarity. It was also observed that quasi-breakdown in thin oxide (20 Å) is distinctly different from that of thicker oxides (45 Å) as shown by the evolution of interface traps and oxide bulk traps at onset of QB. The results suggest that for QB in thin oxide (20Å), a linked conduction path or percolation path is formed while in thicker oxides, damage is confined to either the anode or cathode without forming a linked path. The result shows that the interface trap enhanced tunneling (ITET) is an important mechanism for gate leakage current in thin oxides (< 20 Å) at low to moderate gate voltage ($|V_g|< 3$ V) within the range of normal operating conditions.

As oxide thickness is further scaled downward to its ultimate limit of around 13 Å, breakdown evolves to a progressive characteristics but still retaining its local and discrete nature. It is observed that gate leakage current in ultra-thin oxide 13.4 Å, does not show very significant discrete increases, but increases progressively with localized degradations. This progressive increase in leakage current bears a power relation to stressing time in the initial stage and a linear logarithmic time and voltage dependence in the second PBDs stage. Using multiple breakdown spots statistics, it is shown that the Weibull distribution for the ultra-thin gate oxide can be obtained from a single large area device.

Reference

[6.1] S. H . Lee, B. J. Cho, J.C. Kim, and S.H. Choi, "Quasi-breakdown of ultra thin gate oxide under high field stress," *IEDM Tech. Dig.,* pp 605-608, 1994.

[6.2] Z. Xu, B. J. Cho, and M. F. Li, "Annealing behavior of gate oxide leakage current after quasi-breakdown," *Microelectronic Reliability*, vol. 40, p.1341, 2000.

[6.3] W. Y. Loh, B. J. Cho, and M. F. Li, "Evolution of quasi-breakdown in thin gate oxides," *J. Appl. Phys*., vol. 91, pp. 5302-5306, 2002.

[6.4] M. Depas, T. Nigam, and M. M. Heyns, "Soft breakdown of ultra-thin gate oxide layers," *IEEE Trans. Electron Devices*, vol. 43, pp. 1499-1503, 1996.

[6.5] K. Okada and K. Taniguchi, "Electrical stress-induced variable range hopping conduction in ultra-thin silicon dioxides," *Appl. Phys. Lett.,* vol. 70, pp. 351-353, 1997.

[6.6] D. Ielmini, A. S. Spinelli, M. A. Rigamonti, and A. L. Lacaita, "Modeling of SILC based on electron and hole tunneling. I. Transient effects," *IEEE Trans. Electron Dev.,* vol. 47, pp. 1258-1265, 2000.

[6.7] D. Ielmini, A. S. Spinelli, M. A. Rigamonti, and A. L. Lacaita, "Modeling of SILC based on electron and hole tunneling. II. Steady-state," *IEEE Trans. Electron Devices*, vol. 47, pp. 1266-1272, 2000.

[6.8] J. Maserjian and N. Zamani, J, "Observation of positively charged state generation near the Si/SiO_2 interface during Fowler-Nordheim tunneling," *J. Vac Sci Technol*., vol. 20, pp. 743-746, 1982.

[6.9] P. P. Apte and K. C. Saraswat, **"**Correlation of trap generation to charge-to-breakdown (Q_{bd}): a physical-damage model of dielectric breakdown," *IEEE Trans. Electron Dev.,* vol**.** 41, pp. 1595-1602, 1994.

[6.10] R. Moazzami and C. Hu, "Stress-induced current in thin silicon dioxide films," *IEDM Tech. Dig.,* pp. 139-142, 1992.

[6.11] H. Guan, B. J. Cho, M. F. Li, Z. Xu, Y. D. He, and Z. Dong, "Experimental evidence of interface-controlled mechanism of quasi-breakdown in ultra-thin gate oxide," *IEEE Trans. on Electron Devices*, vol. 48, No. 5, p. 1010, May 2001..

[6.12] P. E. Nicollian, M. Rodder, D. T. Grider, P. Chen, R. M. Wallace, and S. V. Hattangady, "Low voltage stress induced leakage current in ultrathin gate oxides," *37th Proc. Int. Reliab. Phys. Symp*., San Diego, pp. 400-404, 1999.

[6.13] A. Ghetti, E. Sangiorgi, J. Bude, T. W. Sorsch, and G. Weber, "Low Voltage Tunneling in Ultra-thin Oxides : A Monitor for Interface States and Degradation," *IEDM Tech Dig.*, p. 731, 1999.

[6.14] F. Crupi, C. Ciofi, A. Germano, G. Iannaccone, J. H. Stathis, and S. Lomnardo, "On the role of interface states in low-voltage leakage currents of metal-oxide semiconductor structures," *Appl. Phys. Lett*., vol. 80, pp. 4597-4599, 2002.

[6.15] A. Meinertzhagen, C. Petit, D. Zander, O Simonetti, T. Maurel, and M. Jourdain, "Low voltage stress induced leakage current and surface state in ultra-thin (1.2-2.5 nm) oxides," *J. Appl. Phys*., vol. 91, pp. 2123-2132, 2002.

[6.16] M. Depas, T. Nigam, and M. M. Heyns, "Definition of dielectric breakdown for ultra thin (< 2nm) gate oxides," *Solid State Electronics*, vol. 41, pp. 725-728, 1997.

[6.17] J. Sune, E. Y. Wu, D. Jimenez, R. P. Vollertsen, and E. Miranda, " Understanding Soft and Hard Breakdown statistics, prevalence ratios and energy dissipation during breakdown runaway," *IEDM Tech. Dig,* pp. 117-120, 2001.

[6.18] M. A. Alam, B. Weir, J. Bude, P. Silverman, and D. Monroe, " Explanation of soft and hardbreakdown and its consequences for area scaling," *IEDM Tech. Dig,* pp. 449-452, 1999.

[6.19] B. E. Weir, P. J. Silverman, D. Monroe, K. Krisch, M. A. Alam, G. B. Alers, T. W. Sorsch, G. L. Timp, F. Baumann, C. T. Liu, Y. Ma, and D.

Hwang, "Ultra-thin gate dielectrics : They break down but do they fail ?," *IEDM Tech Dig.*, pp 73-76, 1997.

[6.20] Y. Wu, D. Bang, G. Lucovsky, and M. R. Lin, "Time Dependent Dielectric Wearout (TDDW) Technique for Relaibility of Ultrathin Gate Oxides," *IEEE Electron Dev. Lett.,* vol. 20, no. 6, pp. 262 – 264, 1999.

[6.21] F. Monsieur, E. Vincent, D. Roy, S. Bruyere, J. C. Vildeuil, G. Pananakakis, and G. Ghibaudo, "A Thorough Investigation of Progressive Breakdown in Ultra-thin Oxides. Physical Understanding and Application for Industrial Reliability Assessment," *IEEE 40th Int. Reliability Phys. Symp. (IRPS),* pp. 45-54, 2002.

[6.22] E. Farres, M. Nafria, J. Sune, and X. Aymerich, "The statistical distribution of breakdown from multiple breakdown events in one sample," *J. Phys. D,* vol. 24, pp. 407-414, 1991.

[6.23] J. Sune and E. Wu, "Statistics of successive breakdown events for ultra-thin gate oxides," *IEDM Tech Dig.*, pp 147-150, 2002.

[6.24] M. Alam and R. K. Smith, "A Phenomenological theory of correlated multiple soft-breakdown events in ultra-thin gate dielectrics," *IEEE 41th Int. Reliability Phys. Symp. (IRPS),* pp. 406-411, 2003.

[6.25] M. Alam, R. K. Smith, B. Weir, and P. Silverman, "Statistically independent soft-breakdown redefine oxide reliability specifications," *IEDM Tech. Dig.,* pp. 151-154, 2002.

[6.26] M. Alam, R. K. Smith, B. Weir, and P. Silverman, "Uncorrelated breakdown of silicon integrated circuits," *Nature*, 6914, p. 378, 2002.

[6.27] D. Ielmini, A. S. Spinelli, A. L. Lacaita, and M, J. van Duuren, "Correlated defect generation in htin oxides and its impact on Flash reliability," *IEDM Tech. Dig.,* pp. 151-154, 2002.

[6.28] F. Monsieur, E. Vincent, D. Roy, S. Bruyere, J. C. Vildeuil, G. Pananakakis, and G. Ghibaudo, "Evidence for defect-generation-driven wear-out of breakdown conduction path in ultra-thin oxides," *IEEE 41th Int. Reliability Phys. Symp. (IRPS),* pp. 424-431, 2003.

[6.29] B. P. Linder, J. H. Stathis, D. J. Frank, S. Lombardo, and A. Vayshenker, "Growth and Scaling of Oxide Conduction after Breakdown," *IEEE 41th Int. Reliability Phys. Symp. (IRPS),* pp. 402-405, 2003.

[6.30] E. Miranda, J. Sune, R. Rodriguez, M. Nafria, and X. Aymerich, "A Functional-Fit Model for the Soft Breakdown Failure Mode," *IEEE Electron Dev. Lett.,* vol. 20, pp. 265-267, 1999.

[6.31] T. Hosoi, P. L. Re, Y. Kamakura, and K. Taniguchi, "A new model of time evolution of gate leakage current after soft breakdown in ultra-thin gate oxides," *IEDM Tech Dig.*, pp. 727-730, 2002.

[6.32] A. Neugroschel, C. T. Sah, K. M. Han, M. S. Carroll, T. Nishida, J.T. Kavalierios, and Y. Lu, "Direct-current measurements of oxide and traps on oxidized silicon," *IEEE Trans. Electron Dev.,* vol. 42, pp.1657-1662, 1995.

[6.33] B.B Jie, M.F. Li, C.L. Lou, K.F. Lo, W.K. Chim, D.S.H. Chan, "Direct-current measurements of interface traps and Oxide traps in LDD pMOSFETs with an n-well structure", *IEEE 6th Proc. IPFA*, pp. 176-180, 1997.

[6.34] C. Leroux, D. Blachier, O. Briere, and G. Reimbold, *Microelectronic Engineering,* vol. 36, p. 297 (1997).

[6.35] E. Vincent, "Ultrathin oxide breakdowns, an overview," *IEEE Int. Integrated Reliability Workshop Final Report*, pp. 158-166, 1999.

[6.36] A. Ghetti, J. Bude, and G. Weber, "T_{BD} prediction from measurement at low field and room temperature using a new estimator," *Symp. VLSI Technology,* p. 218, 2000.

[6.37] L. K. Han, M. Bhat, D. Wrister, J. Fulford, and D.L. Kwong, "Polarity Dependence of delectrics breakdown in scaled SiO_2," *IEDM Tech. Dig*., pp. 617-620, 1994.

[6.38] S. M. Rowland, R. M. Hill, and L.A. Dissado, "Censored Weibull statistics in the dielectric breakdown of thin oxide films," *J. Phys. C*, vol. 19, pp 6263-6285, 1986.

Chapter 7

High-*K* Dielectrics Reliability: Charge Trapping and Breakdown characteristics

7.1 Introduction

According to the ITRS (International Technological Roadmap Semiconductor) 2003 [7.1], the technological limit for using silicon dioxide (SiO_2) will be reached by 2006 for the 70 nm technology node (low power application) due to its excessive leakage current. As a result, high-*K* gate dielectric replacement for the current SiO_2-based dielectrics is mandatory for 65 nm technology node and beyond [7.1]. Before high-*K* gate material can be used commercially, the reliability and lifetime of such materials have to be properly investigated and characterized to assure sufficient lifetime and reliability margin. In this aspect, the reliability of several high-*K* dielectric films such as ZrO_2, Al_2O_3 and HfO_2 has been investigated by many different groups [7.2]-[7.20],[7.24], and this trend has continued unabated. In spite of this, the breakdown mechanism of the high-*K* with metal gate is still poorly understood. In terms of reliability for high-*K* dielectrics, various problems have already been reported. It was observed that high-*K* dielectrics have low Weibull slopes [7.2],[7.3] which are attributed to extrinsic defects or to a larger percolation/defect cell size respectively. Yamaguchi *et al*. further showed that the bimodal breakdown mechanism, which could increase the breakdown distribution, resulted when crystallization of the dielectric film occurred under temperature annealing higher than 1000^oC [7.4]. In other aspects, charge trapping and threshold instability in high-*K* dielectrics have also been reported [7.6],[7.7],[7.9]-[7.17], which can prove to be a potential problem for high-*K* dielectrics integration. Degraeve *et al.* have observed that polarity dependent breakdown occurs either at the interfacial layer or within the bulk high-*K* layer for Metal-Oxide-Semiconductor (MOS) capacitors [7.19]. This conclusion is supported by a high Weibull slope for positive polarity stress and a low thickness independent Weibull slope under negative polarity stress. Their observation

on dual breakdown mode was made in the relatively thick Al_2O_3 (7-20 nm) and ZrO_2 (8-20 nm) on SiO_2 (1.4-3.6 nm) stacks and the interfacial layer breakdown was observed predominately under gate injection. In HfO_2 however, it is observed recently that low Weibull slopes are obtained even for positive polarity stress, indicating that interfacial SiO_2 breakdown is possible in HfO_2/SiO_2 stack [7.20]. The propensity for interfacial layer breakdown in high-*K* stack is attributed to higher field across the interfacial layer [7.18] although it is unclear if this is still true for other high-*K* materials and as high-*K* layer thickness scales downward to the thickness of interest.

Currently, HfO_2 and $HfAl_xO_y$ alloy film have drawn considerable attention due to its high dielectric constant, acceptable band offset and good thermal stability [7.21]-[7.23]. Process and reliability studies on hafnium dioxide (HfO_2) dielectrics have shown that highly reliable gate stack can be manufactured using HfO_2 with TaN as gate electrode [7.24]-[7.26]. Although HfO_2 have superior electrical characteristics, it suffers from relatively low crystallization temperature. In order to improve its crystallization temperature, Al and nitrogen are usually incorporated into the dielectrics film.

In this chapter, the reliability of complementary MOSFETs with $HfAl_xO_y$ gate dielectric and tantalum nitride (TaN) metal gate is studied. New findings on charge trapping characteristics, breakdown behavior, and statistical analysis are reported and analyzed using the carrier separation technique, which can provide direct evidence on the breakdown mechanism in high-*K* dielectrics stacks. The organization of this chapter is as follows: Section 7.2 describes the device fabrication. Section 7.3 describes briefly the carrier separation measurement technique and its application in high-*K* gate dielectrics. Section 7.4 describe some theoretical aspect of polarity dependent charge trapping measurement using both current-voltage and capacitance-voltage measurements. Section 7.5 shows the experimental results, which includes charge trapping and polarity dependent breakdown in high-*K* stack. Statistical breakdown data are also included for high-*K* stack lifetime studies. Section 7.6 describes the proposed model based on the experimental evidences presented in the previous sections and finally, section 7.7 summarizes the main findings for high-*K* reliability using a novel analytical method based on carrier separation measurements.

7.2 Device Fabrication

The p- and n-MOSFETs used in this study were fabricated on n-type and p-type (100) Si substrate (4~8 Ω·cm). After SC1 cleaning with diluted HF dip, surface nitridation of the silicon wafer at 700°C in an NH_3 ambient was performed to form an oxynitride interfacial layer [7.27]. $HfAl_xO_y$ (thereafter as HfAlO) gate dielectric with an equivalent oxide thickness (EOT) of 17~19.5 Å was then deposited using metal organic chemical vapor deposition (MOCVD) method in a multi-chamber cluster tool, followed by post-deposition annealing at 700°C for 1 minute in a N_2 ambient. The HfAlO film was deposited at 450°C using liquid delivery system (LDS) to transport a cocktail source $HfAl(MMP)_2(OiPr)_5$ [7.28]. TaN gate electrode was then deposited using reactive sputtering with both argon and nitrogen flow. Rapid thermal annealing (RTA) for source/drain activation was performed at 900°C, followed by forming gas annealing at 420°C. The EOT was extracted by UC Berkeley *C-V* simulation program [7.29], taking into account quantum-mechanical effect.

7.3 Carrier Separation and Leakage Path Mechanism

The main difference between conventional silicon dioxide and high-*K* dielectrics in terms of leakage path is the existence of an interfacial layer (IL) and different valence and conduction band offset of the high-*K* stack. Fig. 7.1 shows the energy band diagram of the high-*K* stack of p-MOSFET under both inversion and accumulation conditions. From Fig. 7.1(a), for p-MOSFET, under negative V_g, electron from gate is injected through the high-*K* and interfacial SiO_2 layer and collected at the n-substrate. Conversely, holes are injected from source side and are channeled to the gate through the SiO_2 and part of the high-*K* stack. The main difference between the electron and hole leakage path is the difference in leakage path through the high-*K* layer. Due to the lower permittivity of the IL, significant voltage drop occurs across the IL. As a result, the hole current J_S tunnels mainly through the IL layer but through a smaller portion of the high-*K* layer. In contrast, electron current, J_W, tunnels through both IL and bulk high-*K* layer. In the inversion bias regime (negative gate bias), the hole current as reflected by J_s, is dominant over the electron current J_w due to the usage of TaN as gate electrode which has a midgap

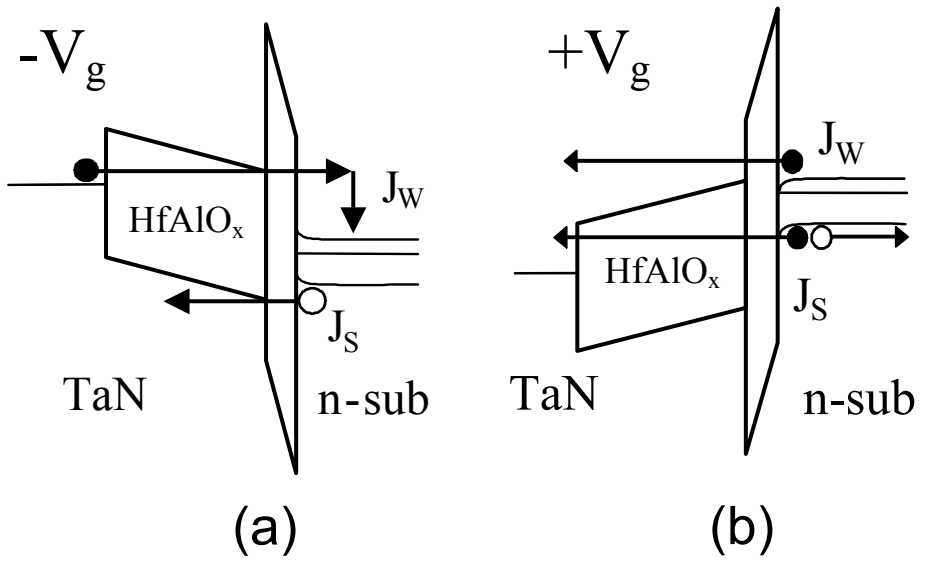


Fig. 7.1 Energy band diagram and tunneling current components for p-MOSFET with metal gate under (a) inversion (negative gate bias) and (b) accumulation (positive gate bias) conditions. The dominant components of gate currents under both polarities are the ones which tunnel through the IL: J_s under $-V_g$ and J_w under $+V_g$.

workfunction of about 4.5-4.7 eV [7.32]. Hole dominance over electron using metal gate electrode for p-MOSFETs under negative gate bias is verified by simulation using modified WKB approximation for hole and electron quantum-mechanical direct tunneling and has been demonstrated previously by Hou *et al* [7.33]. The significance of this method lies in monitoring the difference in leakage path through the bulk high-*K* layer. Under accumulation condition, $V_g > 0$ V (Refer to Fig. 7.1(b)), only electron injection from substrate occurs, due to lack of free holes from the metal gate. For p-MOSFETs, under + V_g, J_w shows the electron current mainly through the IL layer while J_S shows the electron injection through both the IL and bulk high-*K* layer.

In the case of n-MOSFET, the scenario is more complicated under $-V_g$ due to the accumulation of holes in the p-substrate which may recombine with the injected electrons from the gate. However, assuming unity recombination, this method should still be applicable for $-V_g$ in n-MOSFETs. Due to charge recombination at the substrate under accumulation condition, n-MOSFETs data are included in the later part of this chapter for the positive gate bias analysis to eliminate any errors introduced due to charge recombination effects.

As shown in Fig. 7.1, the dominant leakage current shows the leakage path through the interfacial layer and a smaller portion of the high-*K* layer. In contrast, the lower leakage current shows the leakage path through the IL and a larger portion of the high-*K* layer. By monitoring the changes in the dominant and subservient leakage current, the state of the high-*K* layer and IL can be deduced. Fig. 7.2 shows the carrier separation measurement for p-MOSFET under both positive and negative gate bias. As can be inferred from the carrier separation measurement, the tunneling component through the IL is much higher than that through the high-*K* and IL stack. The result is similar to that of M. Houssa et al. using a $SiON/ZrO_2$ stack but without minority holes due to the usage of capacitor structures [7.34].

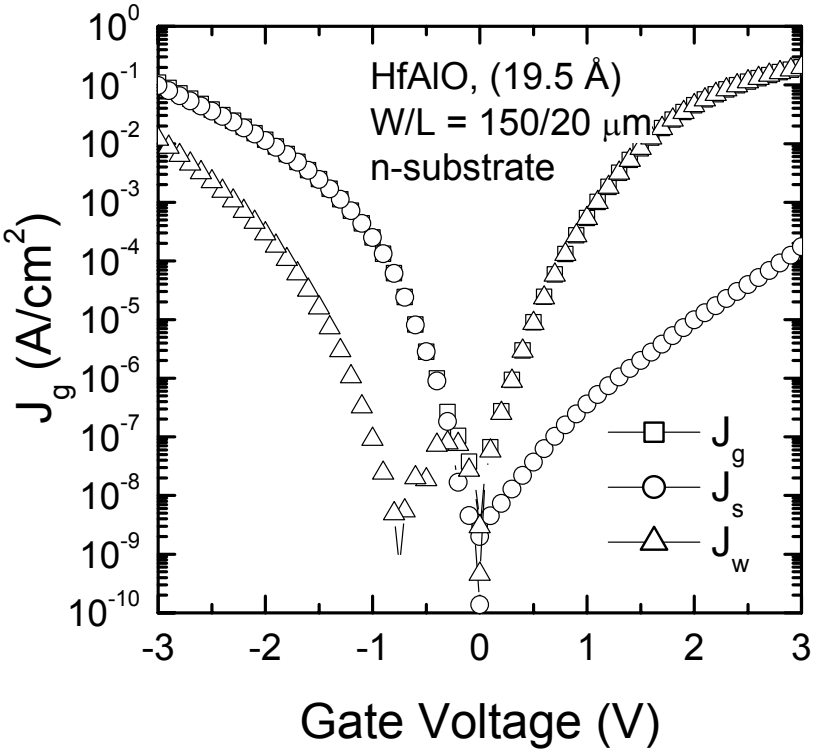


Fig. 7.2 Jg-V_g characteristics of p-MOSFETs with HfAlO dielectrics. Source and n-well are grounded. The dominant components of gate currents under both polarities are the ones which tunnel through the IL: J_s under – V_g and J_w under +V_g.

7.4 Polarity Dependent Charge Trapping

Charge trapping in high-*K* dielectric stack is studied using both current-voltage (*I-V*) and capacitance-voltage (*C-V*) sensing techniques [7.6]. When a dielectric is subjected to a constant voltage stress, electron and hole trapping occurs which changes the charge distribution within the dielectric. The change in flatband voltage ΔV_{FB} under a constant voltage stress due to trapped charges is governed by the following equation (7.1) which relates the change in flatband voltage to the charge trapping within the high-*K* stack.

$$\begin{aligned}\Delta V_{FB} &= -\frac{q}{C_{ox}}\int_{t_{ox}}^{0}\frac{xN^{+}(x)}{t_{ox}}dx \\ &= -\frac{qN}{C_{ox}}(1-\frac{x}{t_{ox}}) \\ &= -\frac{1}{C_{ox}}Q\ (1-\bar{x})\end{aligned} \tag{7.1}$$

where Q is the equivalent trapped charge density and charge centroid $\bar{x}$ is given by the normalized distance x/t_{ox} from the Si substrate interface. From Eq. (7.1), it can be observed that a smaller charge centroid (trapped charges nearer to Si substrate) will have a larger effect on the flatband voltage shift.

In contrast, for *I-V* characteristics, the effect of trapped charges nearer to the injection side has a larger effect than that of those closer to the anode. This can be inferred from the charge centroid relation to that of the voltage change under a ramped voltage stress. From Fischetti *et al*. [7.35] and Solomon *et al.* calculation, assuming a positive charge sheet N^{+} charges/cm^2 located at a distance of *x* from the Si interface, it has been shown that under a negative gate injection :

$$\begin{aligned}\Delta V_{IV} &= -\frac{qN^{+}t_{ox}}{\varepsilon_{ox}}\left(\frac{x}{t_{ox}}\right) \\ &= -\frac{Q^{+}}{C_{ox}}.(\bar{x})\end{aligned} \tag{7.2}$$

where ε_{ox} is the dielectric permittivity and *q* is electronic charge, $\bar{x}$ is the normalized distance x/t_{ox} from the SiO_2-Si interface and Q is the total trapped charges.

Using the combined ΔV_{CV} and ΔV_{IV}, the trapped charge centroid can be deduced and calculated. This was given in [7.6] as

$$\bar{x} = \frac{1}{1 + \frac{\Delta V_{CV}}{\Delta V_{IL}}} \tag{7.3}$$

Hence, by comparing the voltage change in *C-V* and *I-V* measurement, a rough estimate of the positive charge centroid can be deduced.

7.5 Experimental Results

7.5.1 Charge Trapping in High-*K* stacks

Charge trapping characteristics during constant voltage stressing (CVS) under both positive and negative bias polarities on p-channel (n-substrate) MOS capacitors are investigated. Under positive bias stressing, a strong negative charge trapping is observed, while under negative bias stressing, positive charge or hole trapping is observed for $|V_g| > 3$ V as shown in Fig. 7.3(a) and (b) respectively. This positive charge trapping characteristic is also confirmed by the flat-band voltage shift in C-V curve during negative bias stressing in Fig. 7.4(a). However, under positive bias stressing (negative charge trapping), the C-V curve shows negligible shift in V_{FB} as shown in Fig. 7.4(b). It should be noted that the gate leakage current under a CVS is more sensitive to charge trapping near to the injection side, while the C-V measurement senses the trapped charges closer to the SiO_2/Si interface as shown in Eqns. (7.1) and (7.2). The C-V curve shift behaviors imply that negative charge trapping occurs far from the substrate while positive charge trapping occurs nearer to the HfAlO/SiO_xN_y interfacial oxynitride layer (IL) hence having a larger effect on V_{FB} shift. For p-channel MOSFETs under high negative gate bias, both electrons and holes are injected from the metal electrode and n-substrate respectively. In contrast, under positive gate bias stress, only electron injection occurs due to the lack of free holes in TaN electrode. The origin of the positive and negative charge trapping observed under gate and substrate injection respectively has not been fully verified.

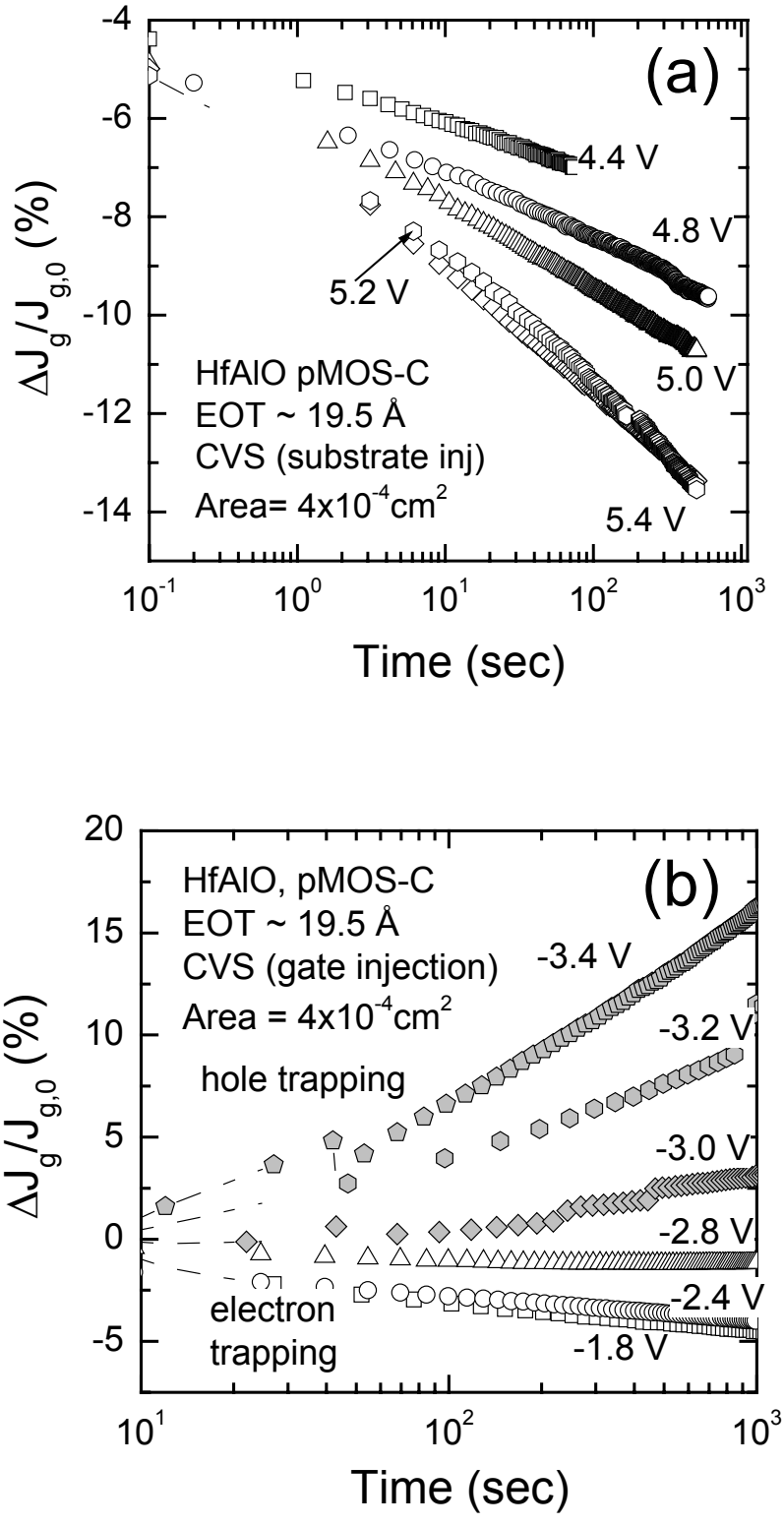


Fig. 7.3 Charge trapping characteristics (a) under positive V_g, and (b) negative V_g. Positive bias stress causes electron trapping, while negative bias stress results in hole trapping for $|V_g| > -3$ V. Sample area is 100 x 100 μm^2.

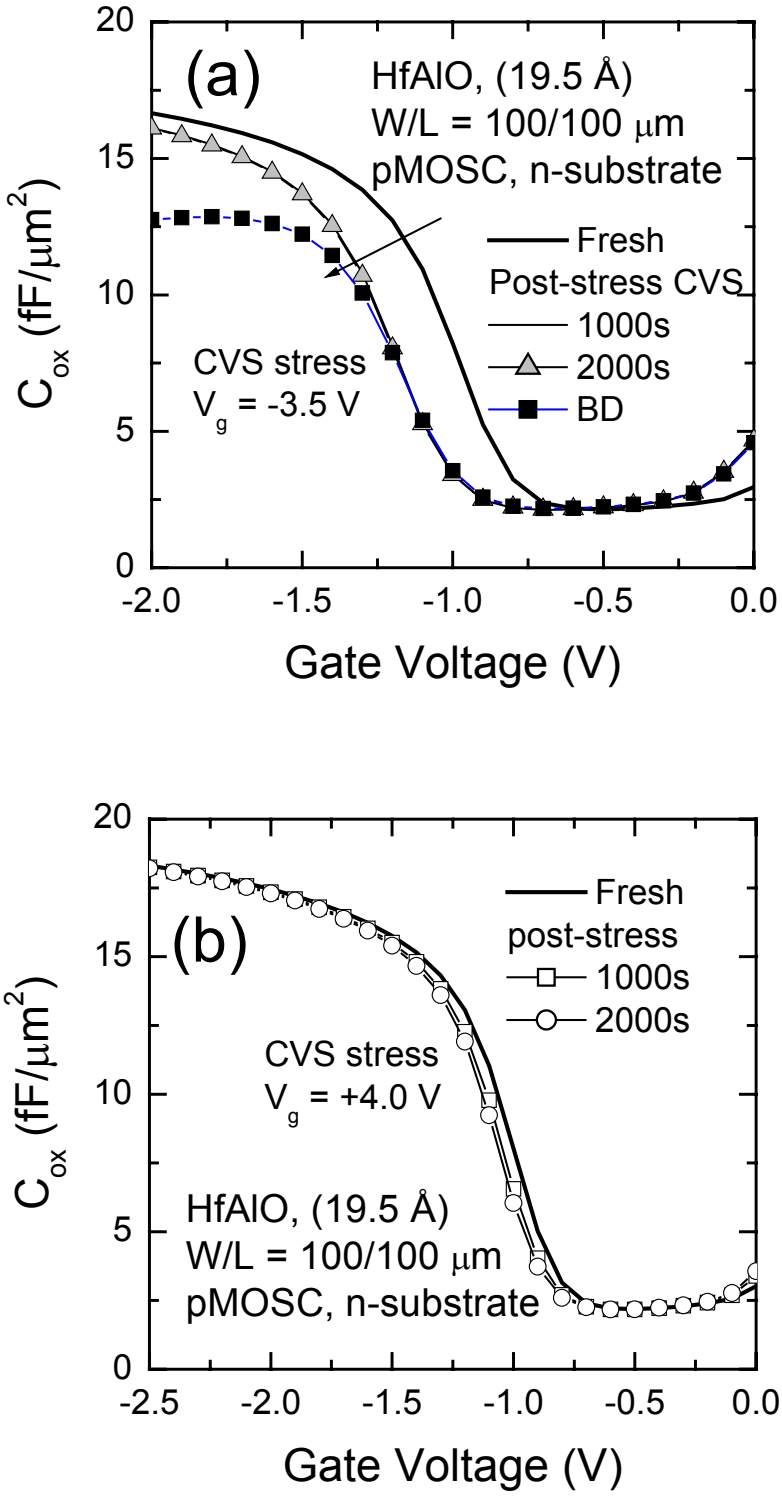


Fig. 7.4 Inversion capacitance-voltage curves of p-MOSFETs before and after stress under (a) negative gate bias and (b) positive gate bias for a period of 1000 s and 2000 s. A flatband voltage shift to the left after negative bias stress shows positive charge trapping within the dielectric and IL.

Using similar I-V and C-V sensing measurement technique, Xu *et al*. have proposed that the positive charge can be ascribed to holes generated by electron impact ionization at the anode which are subsequently trapped in near the SiO_2/HfO_2 interface [7.7]. Another possibility raised is the generation of $[Hf_2{=}OH]^+$ or $[Si_2{=}OH]^+$ centers due to the trapping of H^+ which are released upon electron impact at the anode. [7.7]. However, this possibility is highly unlikely based on the ESR measurements performed by Kang *et al.* which showed that HfO_2/Si has very robust interface without any observed thermodynamic instability as in conventional SiO_2/Si interface [7.8]. Using temperature-current measurement, Blomme *et al.* have further shown that the change in gate leakage current for ZrO_2 under positive gate bias (substrate injection) can be directly attributed to charge trapping which modifies the tunneling barrier [7.9]. Similarly, Xu *et al*. have shown a weakly temperature dependent leakage current for negative gate bias and a strong temperature dependent leakage current for positive gate bias which they ascribe to Frenkel-Poole (F-P) hopping [7.11]. Although the exact mechanism for positive charge generation has not been verified, it is generally agreed that positive charge tends to be trapped near to the SiO_2/Si interface while negative charges are trapped mainly in the bulk of the high-*K* dielectrics [7.6],[7.7],[7.19]. Our experimental results shown in Fig. 7.3 agree with the experimental results observed elsewhere for SiO_2/Al_2O_3 stack with TiN metal electrode [7.19] and its polarity dependent charge centroid [7.6], and for SiON with TiN metal electrode [7.36]. Since the trapping characteristics for different dielectrics and for different gate electrode is similar, the charge trapping mechanism can be attributed to either a difference in the type of carriers injected or due to the difference in interfacial properties as proposed by [7.36] and may not be a property of the bulk material of the dielectric.

As observed from Fig. 7.3(a) and (b), the charge trapping phenomenon in HfAlO is observed to follow a logarithmic dependence with stressing time rather than an exponential dependence as in the case for SiO_2. Different researchers have studied this phenomenon and differing views are proposed. Shanware *et al.* showed that by assuming a distributed time constant and traps location which increase exponentially with distance of the trap from the substrate, a log-time dependence can be obtained

[7.13]. In contrast, Zafar *et al*. showed that simply using a stretched exponential based on a distributed time constant yield a stretched exponential modeling which fits well for the initial stage of stressing [7.14]. In any case, both modeling results in almost linear log-time dependence in the initial stage as observed here. In the case of the stretched exponential modeling, an eventual saturation in charge trapping is expected due to a fixed charge centroid assumption. In our experiment, it is observed that log-time dependence is maintained till breakdown, which is similar to the conclusion of Shanware *et al.* [7.13].

7.5.2 Polarity dependent breakdown characteristics in MOS capacitors

Figure 7.5(a) shows the evolution of gate leakage current during a typical constant voltage stress (CVS) condition on p-channel MOS capacitors with a gate bias V_g = -4.1 V. The corresponding current-voltage (I-V) curves are shown in Fig. 7.5(b). It can be observed that there are at least two distinct stages in the breakdown. In this case, post-breakdown (pBD) stages, pBD1 and pBD2, share similar I-V characteristics while post-breakdown stage, pBD3 shows entirely different characteristics. The multi-modal breakdown in high-*K*/IL stack dielectrics looks similar to quasi-breakdown (soft-breakdown) observed in pure SiO_2. In contrast, the I-V characteristic of pBD3 shows very high gate leakage, with a similar characteristic to complete (hard) breakdown observed in conventional SiO_2. Besides the large increase in gate leakage current after pBD3, it can be observed that gate current fluctuation also increases after onset of pBD3 while pBD1 and pBD2 stages have relatively smaller gate current fluctuation. However, in the case of positive CVS, only one breakdown stage is observed as shown in Fig. 7.6. Fig. 7.6(a) shows the evolution of gate leakage current density under positive bias CVS. Breakdown is observed after approximately 1800 s of constant voltage stress. At the onset of the breakdown, it can be observed that gate leakage current also shows a large current fluctuation similar to that observed after pBD3 for the negative bias CVS (Refer to Fig. 7.5(a)). By comparing the I-V characteristics, it is easy to infer that the single breakdown stage

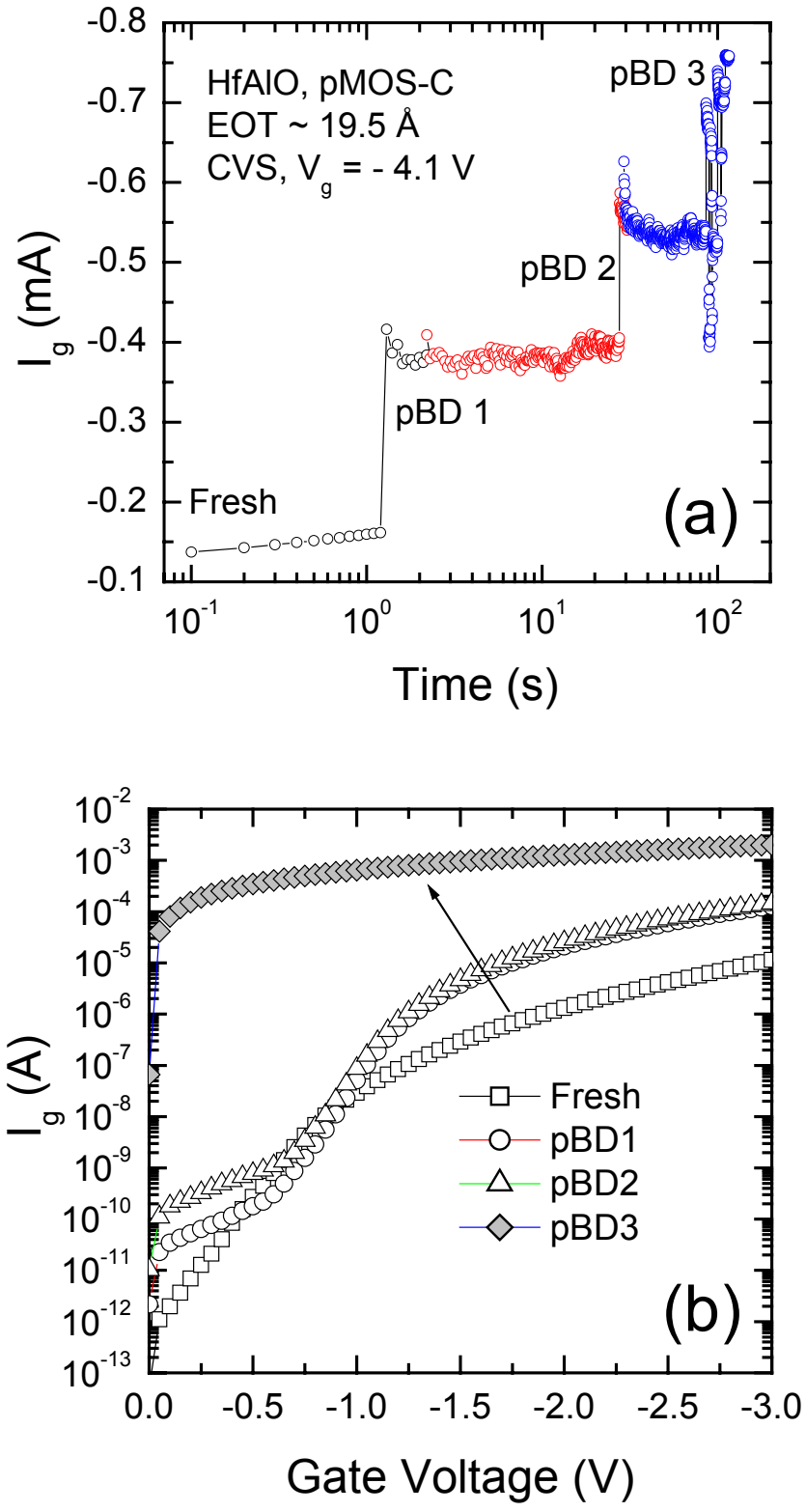


Fig. 7.5 (a) Evolution of gate leakage current in p-MOSC and (b) evolution of current-voltage characteristics under negative constant voltage stress. It can be observed that there are at least 2 distinct stages of breakdown as shown by pBD1, pBD2 and pBD3. Area of sample used is 100 x 100 μm^2.

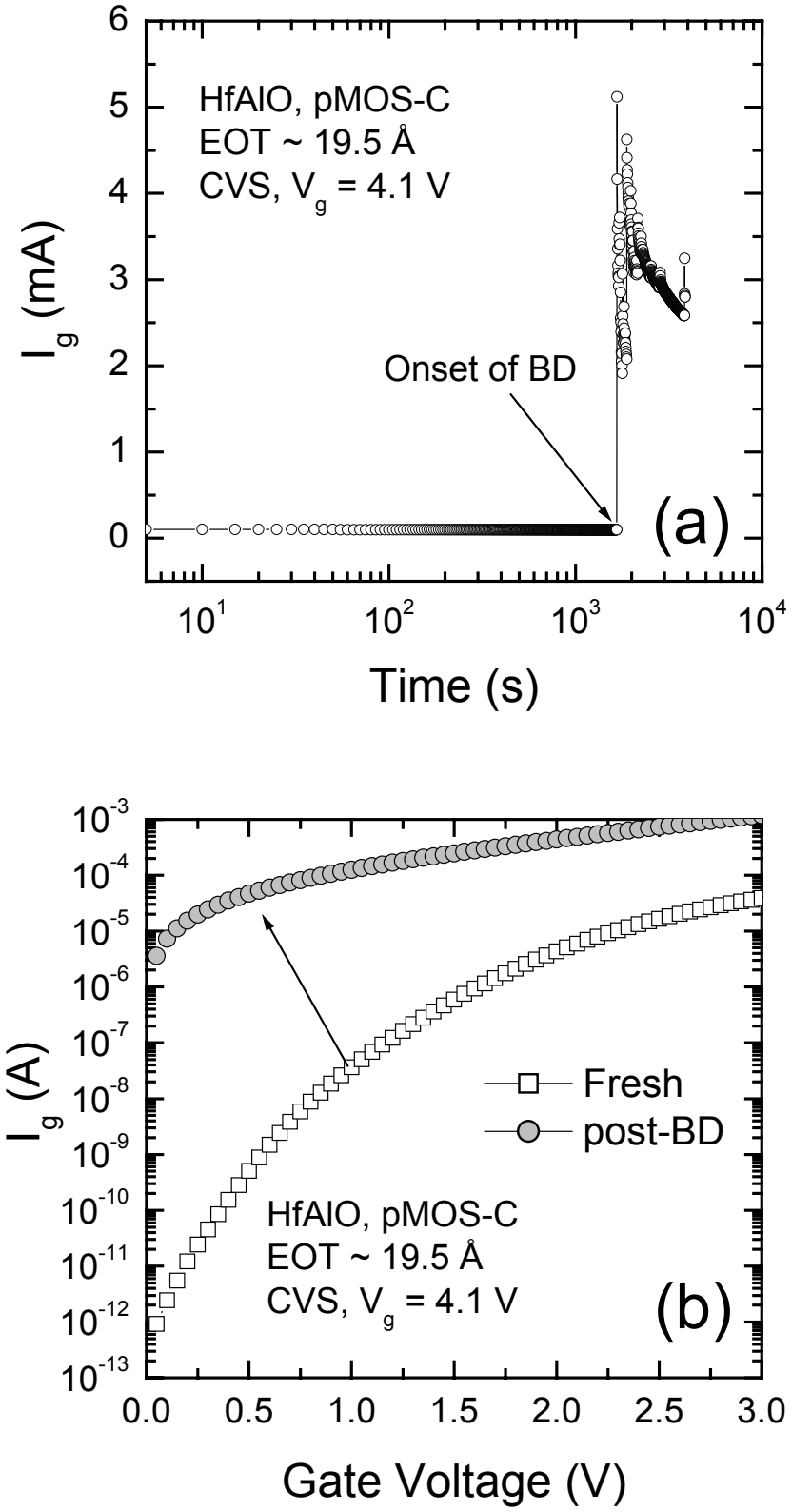


Fig. 7.6 (a) Evolution of gate leakage current in p-MOSC and (b) evolution of current-voltage characteristics under positive constant voltage stress. It can be observed that there are only 1 distinct stages of breakdown. Area of sample used is 100 x 100 μm^2.

observed under positive bias CVS is similar to that of pBD3 under negative bias CVS. However, it is not possible to distinguish the multi-modal breakdown observed under negative bias CVS simply based on I-V characteristics. Moreover, although the I-V characteristics after pBD3 resembles that of complete breakdown in SiO_2, the excessive gate current fluctuation is more similar to that exhibited in oxides after quasi-breakdown. The results suggest that there are a large amount of trapping/detrapping of charges after onset of pBD3 for negative bias CVS and after breakdown for positive bias CVS.

7.5.3 Negative CVS : p-MOSFET

Figure 7.7 shows the relative changes of source (J_s) and well (J_w) currents during negative bias CVS under inversion conditions. The plots of (a) and (b) in Fig. 7.7 are from identical data but plotted in different scales for easy recognition of each component. The result shows that at the initial stage of breakdown under negative gate bias, well current J_w shows a higher percentage change in leakage current than source current J_s. This implies that the high-*K* bulk has broken down first. However, subsequent stressing results in the higher percentage change in source current J_s than J_w, indicating that the IL breakdown becomes the dominant factor at the later stage of breakdown. Fig. 7.8 shows the associated I-V characteristics of the p-channel MOSFET prior and after the breakdown phenomenon. It can be observed that at the onset of Bulk BD, J_s which reflects the IL condition, increases marginally while J_w which reflects the tunneling current through the entire high-*K* stack increases significantly throughout the entire range of the measurement gate voltage. At the onset of IL BD however, both J_s and J_w increase significantly since the bulk layer has broken down and both currents actually reflect the condition of the interfacial layer (IL). The results of both the I-V and carrier separation measurement during CVS for p-MOSFET are consistent and show conclusively that bulk BD occur first under a negative gate bias CVS. Therefore, from the results in Figs. 7.7 and 7.8, it is clear that the multi-modal breakdown phenomenon (pBD1 and pBD2) observed in Fig. 7.5 is related to the degradation in the high-*K* bulk, and not in the IL. It is thus only at the onset of pBD3 in Fig. 7.5 that both the IL and high-*K* bulk have broken down.

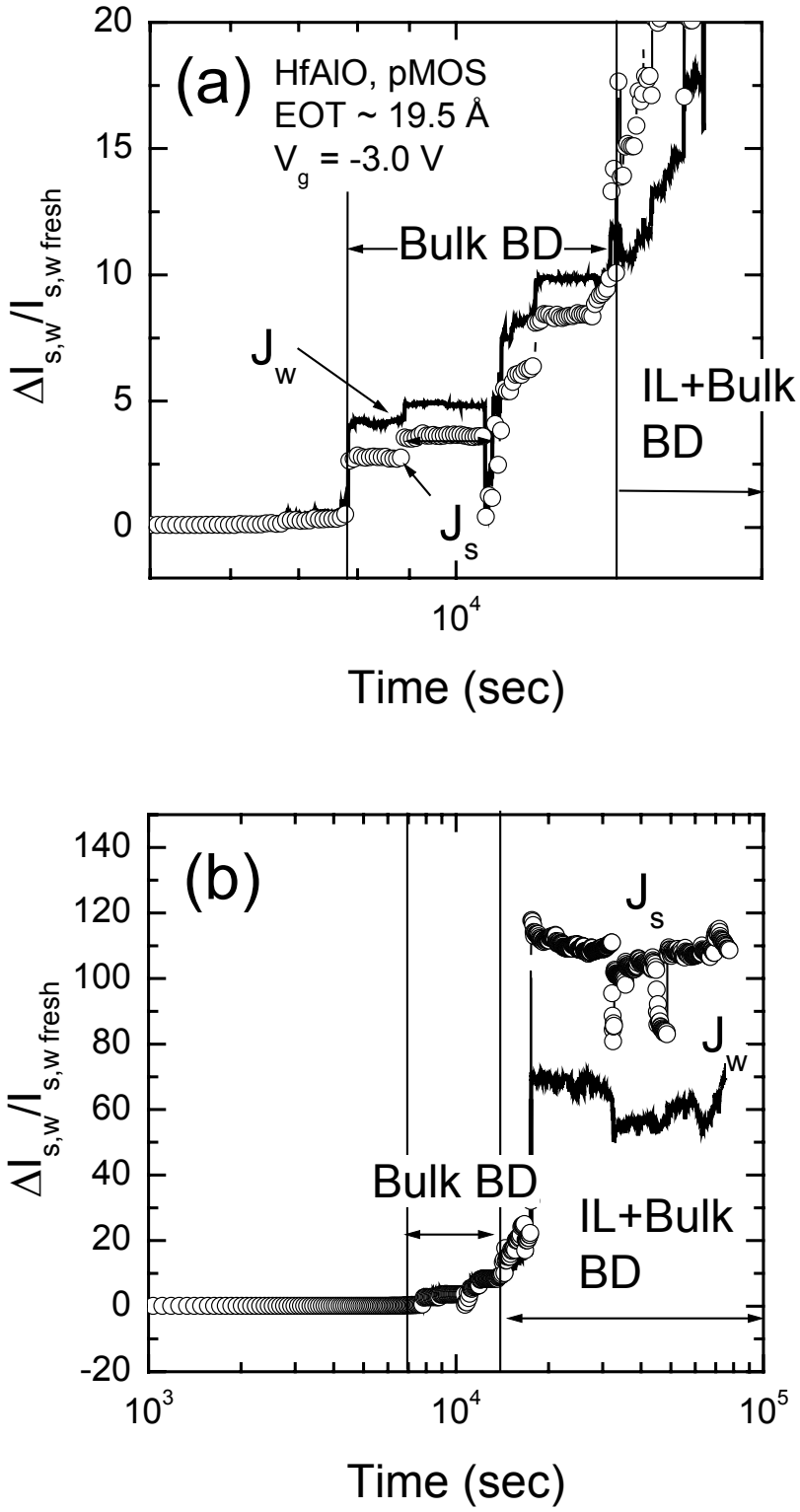


Fig. 7.7 (a) Relative changes of J_s and J_w currents during negative CVS (V_g = -3.0 V) on p-MOSFET. (b) Identical data with (a) but plotted in wider scale. The high-*K* bulk breakdown happens first at the initial stage of breakdown.

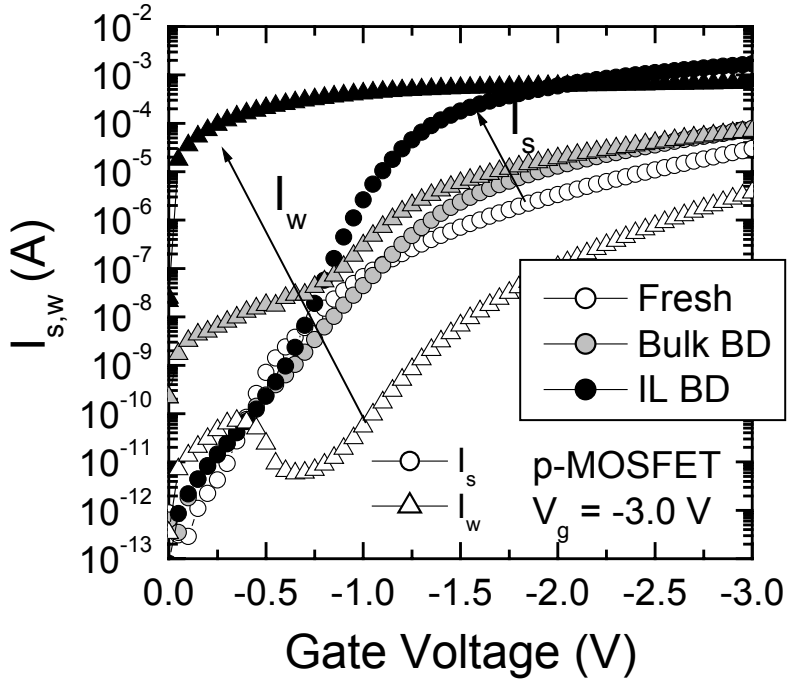


Fig. 7.8 Associated (J-V characteristics) carrier separation measurement of Fig. 6.8, showing both source and well current component before and after bulk BD and IL BD. It can be observed that at the first onset of Bulk BD (shown by shaded symbols), J_w increases significantly throughout the entire range of gate voltage while J_s, which reflect the IL condition, increases only marginally. Subsequent occurrence of IL BD (shown by solid symbols), result in significant increase in both J_s and J_w.

It is interesting to note that at the onset of bulk high-*K* breakdown, bulk current shown by J_w in Fig. 7.8 exhibits step-like increases in current showing that breakdown in bulk high-*K* also show similar localization effect as SiO_2 wearout and breakdown. Moreover, from Fig. 7.8, it can be observed that well current J_w, shows a hump at $|V_g| < 0.7$ V which can be correlated with the alignment to the n-Si substrate forbidden bandgap. At such low voltage, interface trap assisted tunneling (ITAT) is known to dominate. At onset of bulk BD, a large increase in J_w for $|V_g| < 0.7$ V is observed, showing that bulk BD results in enhanced interface trap assisted tunneling (ITAT). This can be explained by the field penetration of the dielectric due to bulk layer breakdown, resulting in the probing of the interfacial state at the anode. Under negative gate bias stress, it has shown that significant interfacial traps are formed at the interface of HfO_2/Si [7.17],[7.19]. When the bulk layer is intact, ITAT is suppressed due to the physical thickness of the bulk layer. However when the bulk

layer breaks down, the gate leakage current will exhibit an enhanced ITAT as it is tunneling mainly through a reduced thickness of dielectric which allows ITAT to exhibit itself. As a result, all of the observed phenomenon for gate injection dielectric breakdown can be successfully explained by the proposed mechanism of a bulk BD followed by IL BD as deduced from the carrier separation measurements results.

Using similar analysis, it was observed from Fig. 7.9(a) that positive CVS results in degradation mainly at the IL layer but almost no degradation in the high-*K* bulk. This is deduced based on the fact that only J_w, which reflects the IL state, have increased while J_s, which reflect the bulk layer condition under positive bias, did not change significantly. Fig. 7.9(b) shows the I-V characteristics after positive bias stress, but measured under both positive and negative gate voltage regime. The IL leakage currents shown by J_s (for - V_g regime) and J_w (for + V_g regime) have increased by orders of magnitude, while the bulk leakage current shown by the lower magnitude J_w and J_s respectively did not increase significantly due to the intact high-*K* bulk layer. This confirms again that the positive bias stress causes the degradation mainly in IL. Comparing Fig. 7.9 (b) and Fig. 7.6(b), we can see that the high leakage current in Fig. 7.9(b) is only due to the increase of J_w (under $+V_g$) which means only that the IL has breakdown, even though the I-V curve in Fig. 7.6(b) can easily be regarded as the breakdown of the entire dielectric stack.

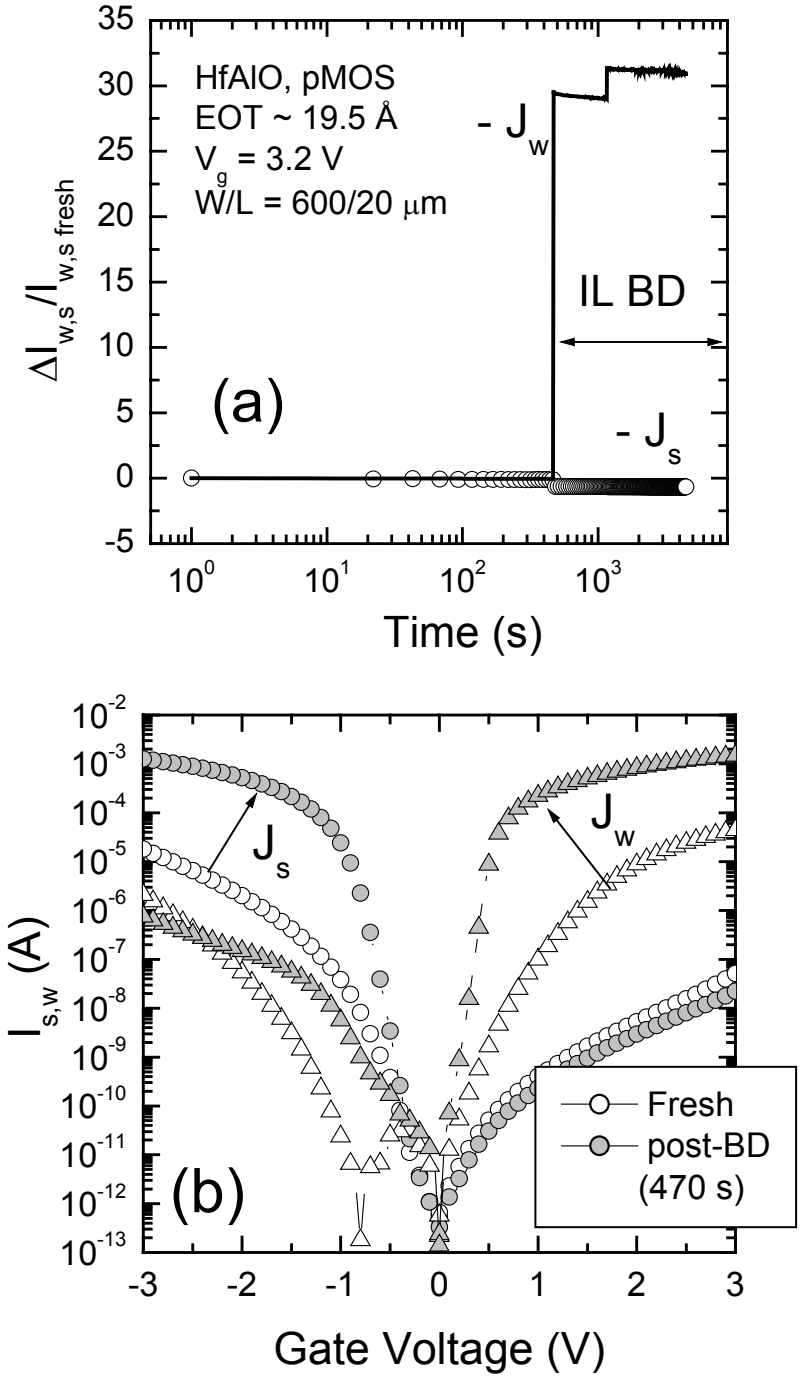


Fig. 7.9 (a) Relative changes of J_s and J_w currents during positive CVS (V_g = +3.2 V) on p-MOSFET. Only interfacial layer breakdown observed after about 470s of stressing, shown by the larger increase in J_w. (b) Associated I-V characteristics after positive CVS but measured under both negative and positive gate voltage regime. The IL leakage currents have increased by orders of magnitude, while the bulk leakage current did not increase significantly due to the intact high-*K* bulk layer.

7.5.4 Positive CVS : n-MOSFET

In order to rule out the effect of electron-hole recombination under accumulation condition, carrier separation measurements are also carried out under inversion conditions on n-channel MOSFET. Figure 7.10(a) and (b) show the carrier separation measurement of n-channel MOSFET stressed under inversion condition at a gate voltage of V_g = 3.2 V. From Fig. 7.10(a), it can be observed that for an n-channel MOSFETs under positive CVS, gate leakage current increases in steps after onset of breakdown, similar to that observed in p-channel MOSFETs and in breakdown of conventional SiO_2 gate dielectrics. Separate monitoring of both J_w and J_s current component shows that at onset of breakdown (after about 15 s of stressing at V_g = 3.2 V), source current J_s is observed to increase proportionally faster than that of well current J_w. I-V characteristics for both J_s and J_w shown in Fig. 7.10(b), show increases in leakage current for both components, indicative of interfacial layer breakdown. This is similar to the conclusion for positive gate stress on p-channel MOSFET, showing that the polarity breakdown of high-*K* stack is substrate independent. We investigated all the possible combinations of polarity dependence in both n- and p- MOSFETs, and the criteria for determining the dominant breakdown mechanism is summarized in Table 7.1.

Table 7.1 A summary of criteria for determining the dominant breakdown mechanism for all the possible combinations of gate bias in both n- and p- MOSFETs

	measurement condition	p-MOSFET	n-MOSFET
IL BD	+Vg	$\Delta Jw > \Delta Js$	$\Delta Js > \Delta Jw$
	-Vg	$\Delta Js > \Delta Jw$	$\Delta Jw > \Delta Js$
Bulk BD	+Vg	$\Delta Jw \leq \Delta Js$	$\Delta Jw \geq \Delta Js$
	-Vg	$\Delta Jw \geq \Delta Js$	$\Delta Jw \leq \Delta Js$

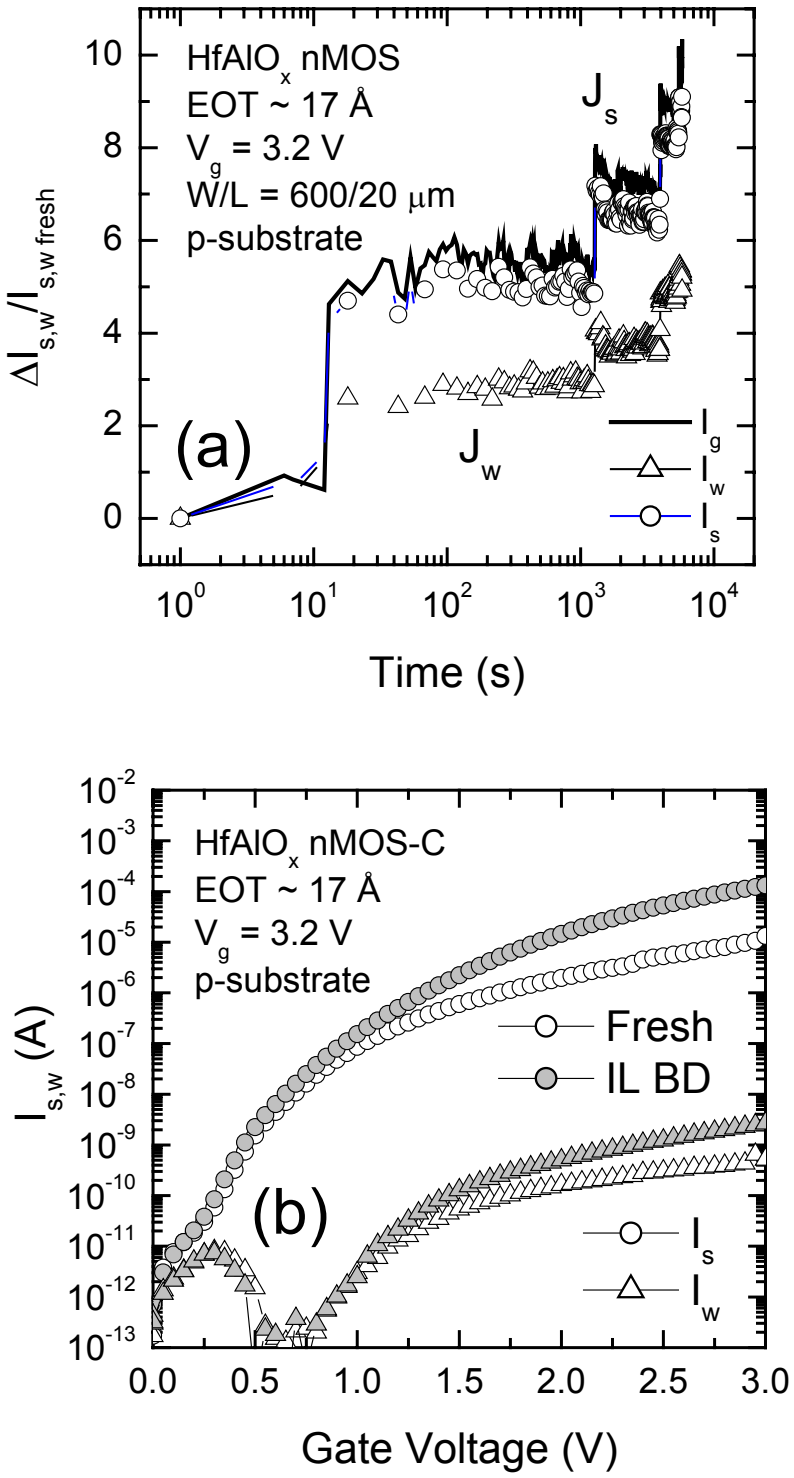


Fig. 7.10 (a) Relative changes of J_s and J_w currents during positive CVS (V_g = 3.2 V) on n-channel MOSFET and (b) its associated J-V characteristics. The dominant electron current, J_s, increase faster than that of the subservient well current, J_w, which reflect the valence electron current. At onset of breakdown after stressing for 15 s, interfacial layer breaks down, which translate to much higher leakage current for J_s while the well leakage current J_w which tunnel through a thicker portion of the high-*K* layer is less affected.

Based on the analysis as shown in Table 7.1, carrier separation measurement results for positive gate bias on n-channel MOSFET (inversion condition) and current component monitoring on p-channel MOSFET(accumulation condition) clearly show that IL has broken down first. In contrast, carrier separation measurement on p-channel MOSFET under negative gate bias (inversion condition) at onset of breakdown shows a bulk initiated breakdown. From the thorough study using carrier separation measurement technique for at least 10 samples each on both p-channel and n-channel MOSFETs under inversion conditions, it was observed that bulk high-*K* layers breakdown occur mainly under negative CVS while IL breakdown occurs mainly under positive CVS for both n- and p-MOSFETs. This is distinctly different from the conclusions for other high-*K* films as previously reported [7.19] but agree well with the recent results obtained using HfO_2 [7.20].

7.5.5 Statistical Breakdown Results

Figure 7.11 shows the Weibull distributions for breakdown voltage (V_{BD}) under ramped gate voltage sweep with a ramp rate of 1V/s for both p-channel and n-channel MOS capacitors. It can be observed that the Weibull slope for negative gate voltage is much steeper than that for positive gate sweep for both n-channel and p-channel MOSFET. Since breakdown is due to a critical level of defect based on percolation or interface-damage theory, the bulk layer which has a larger physical thickness will have a smaller spread in the breakdown distribution than that of the IL. The experimental results of steeper slope for V_{BD} distribution under negative gate bias compared to positive gate bias agree with our initial postulation that bulk BD occurs under negative gate bias and IL BD occurs under positive gate bias. On the other hand, it is also observed that the mean V_{BD} for negative gate bias is smaller than that of positive gate bias as shown in Fig. 7.11. Since IL is physically thinner than bulk, the breakdown voltage should theoretically be much smaller. However, the actual gate bias is applied to both the bulk layer and IL. Hence the voltage drop across interfacial layer, V_{IL} is not equal to the gate bias but is given by the simple voltage division rule as shown in the next page.

$$V_{IL} = \left(V_g - V_{FB}\right)\frac{EOT_{IL}}{EOT_{stack}}$$
$$V_{Bulk} = \left(V_g - V_{FB}\right)\frac{EOT_{bulk}}{EOT_{stack}} \qquad (7.5)$$
$$V_{IL} = V_{Bulk}\left(\frac{EOT_{IL}}{EOT_{Bulk}}\right)$$

Due to this voltage division, the actual voltage drop across IL is a fraction of the applied gate voltage. From [7.3], it has been shown that high-*K* breakdown field for HfO_2 or HfAlO will be smaller of about 4 - 6 MV/cm compared to 12-15 MV/cm for SiO_2 and SiON. Based on a physical thickness of about 12 ~15 Å and 45 ~50 Å for IL and bulk HfAlO layer respectively, the maximum breakdown gate voltage for the SiO_xN_y IL and bulk high-*K* layer works out to be 4.2 V ~ 5.3 V and 3.75 V ~ 4.6 V respectively. The breakdown voltage values obtained experimentally, 5.5 V and 4.5

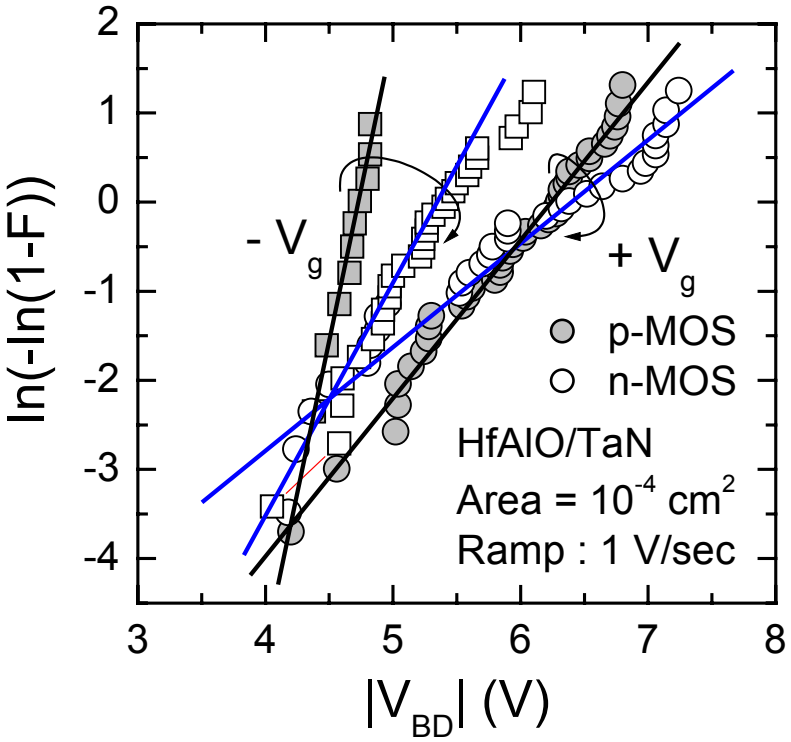


Fig. 7.11 Weibull distributions for V_{BD} under ramped gate voltage sweep for p-channel MOS capacitors (shown by shaded symbols) and n-channel MOS capacitors (shown by open symbols) with source/drain implant. The Weibull slope for breakdown voltage VBD under $-V_g$ sweep is much steeper than that for $+V_g$ sweep for both n-channel and p-channel MOS.

V for IL and bulk layer respectively, corresponds quite closely to the upper limit of the calculated values.

Similar spread is also observed in Charge-to-breakdown (Q_{BD}) distribution measured using constant voltage stress. Q_{BD} distributions for p-channel MOS capacitors as shown in Fig. 7.12, also show reasonably high β (Weibull slope) value under $-V_g$ and low β value under $+V_g$. An interesting phenomenon found in Fig. 7.12 is that the Q_{BD} distribution for negative bias CVS shows voltage dependence while the Q_{BD} distribution for positive bias does not. The negligible voltage dependence for positive CVS (at least at high gate voltage range used in this experiment), implies that IL breakdown is fluency-driven breakdown while bulk high-*K* breakdown which occurs predominantly under negative CVS shows considerable electric field-driven breakdown. This fluency-driven phenomenon for interface layer breakdown is also observed by other researchers [7.19] and has beneficial implication as device scales downward due to voltage scaling which will lead to reduced gate leakage/fluency and hence, lifetime enhancement for IL initiated breakdown.

Figure 7.13 shows the β values plotted against stress voltage and it can be seen that the β values have 2 distinct bands. Under negative bias CVS, the Weibull slope parameter ranges from 1.5 ~ 2.1 which is comparable to the values obtained by other researchers for the same thickness (50 ~ 60 Å) of high-*K* dielectrics [7.12], [7.37]. In the positive CVS regime, however, the β values are extremely low $\beta < 1$ probably due to the ultra-thin SiO_xN_y interfacial layer. The latter may be intrinsically defective in the process, due to its inherent thickness non-uniformity. At low charge fluency, it is observed that Weibull distribution for positive CVS deviates from the linear line fitting. This is especially so for higher CVS and is ascribed to censoring effect arising from resolution of the measurement setup which may have problems in detecting the early breakdowns. In spite of the deviation, it can be observed that Weibull distribution for lower CVS ($V_g = 4.0$ V) can be fitted very well linearly with $\beta = 0.72$. The low Weibull slope under positive CVS is related to the IL quality. Another possibility is a bi-modal degradation effect resulting from a combination of IL initiated breakdown and bulk layer breakdown. Without additional process optimization and conditions splits, it is difficult to determine the exact root cause of

the low Weibull slope under positive gate bias stress for both n- and p-channel MOSFETs. Nevertheless, the results does not detract from the fact that it fits well with the earlier conclusion that IL breakdown is dominant under positive bias stress while bulk high-*K* breakdown is dominant under negative bias stress analyzed using the carrier separation measurement technique.

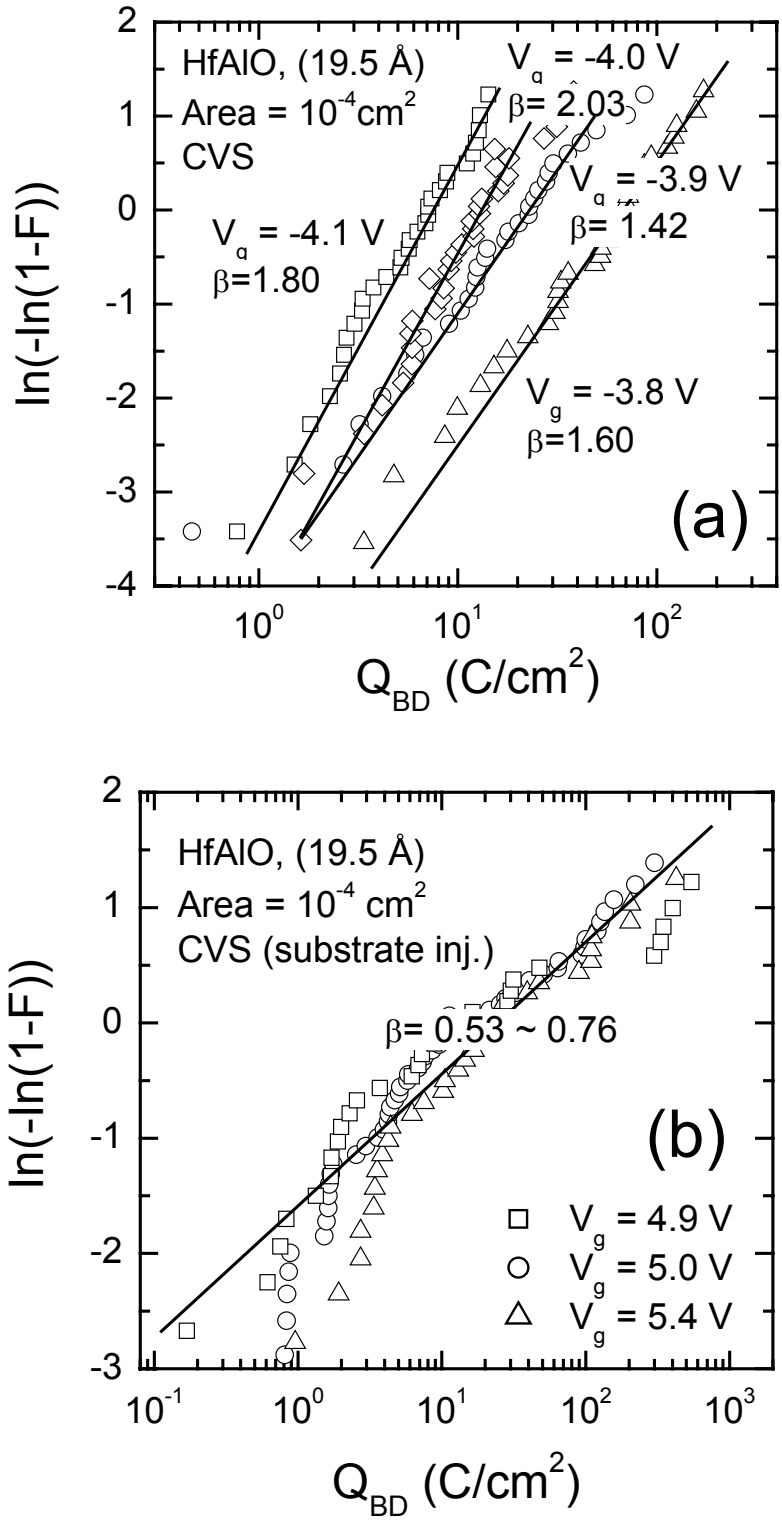


Fig. 7.12 Weibull distributions for charge-to-breakdown (Q_{BD}) (a) under negative constant voltage stresses and (b) under positive constant voltage stresses. It is observed that Weibull distribution for $+V_g$ CVS deviates from the linear line at low Q_{BD} for higher gate bias stressing due to temporal resolution of measurement setup, especially for early failure devices with low time-to-breakdown. High β (Weibull slope) value under $-V_g$ and low β value under $+V_g$ are observed. Sample area is 100 x 100 μm^2.

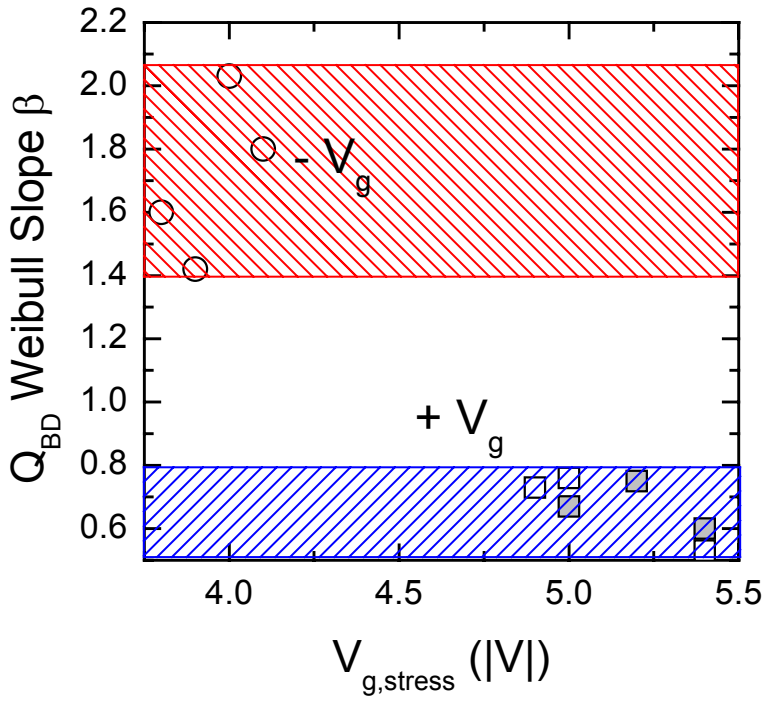


Fig. 7.13 Weibull distributions for charge-to-breakdown (Q_{BD}) on p-channel MOS capacitors under (a) negative constant voltage stresses and (b) under positive constant voltage stresses. Reasonably high β (Weibull slope) value under $-V_g$ CVS and low β value under $+V_g$ CVS are observed. Sample area is 10^{-4} cm^2.

7.6 Proposed Charge Induced Breakdown Model

Based on all the observations, a breakdown model using charge trappings at different spatial locations in high-*K*/IL stack dielectric with a metal gate structure is proposed as shown in Fig. 7.14. Under negative bias CVS, electrons from cathode are injected into the high-*K* dielectrics while holes from n-Si substrate are injected into the high-*K* stack. Electron trapping within the bulk of the HfAlO stack and hole trapping near the IL occur (Fig. 7.14 (a)) as evidenced in charge trapping characteristics and C-V curve shifts in Fig. 7.4. The electron and hole trapping within the HfAlO stack causes distorted band bending and enhances the internal field of the bulk high-*K* material, leading to higher probability of bulk- initiated breakdown. Conversely, for positive gate bias CVS, electron injection from the n-Si substrate will result in the generation of both electron and hole traps in the dielectric. However, due to the lack of free holes in TaN, only electrons trapping will occur. Thus, the band distortion (in the bulk high-*K*) caused by the trapped electrons alone would be smaller

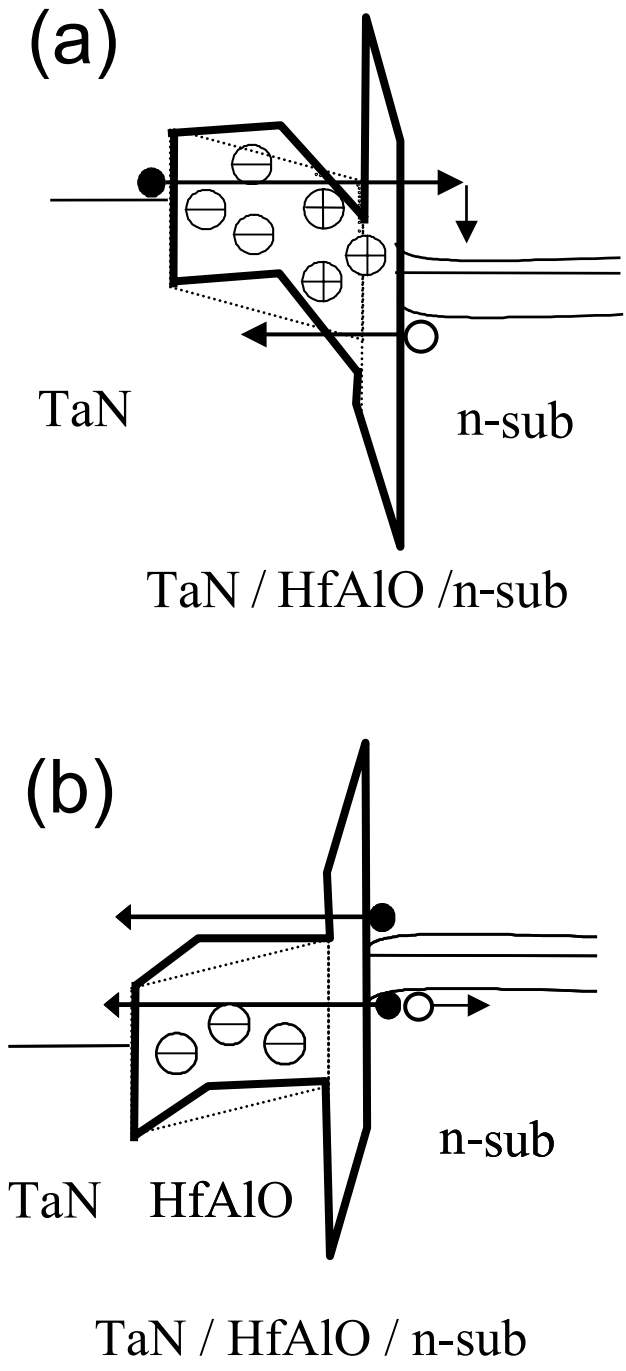


Fig. 7.14 A schematic drawing for a breakdown model using charge trappings at different spatial locations in high-*K*/IL stack dielectric with a metal gate structure. (a) For -Vg, electron trapping occurs mainly in the bulk while hole trapping occurs near to the IL. The columbic force of the trapped charges distorts the energy band diagram, leading to a preferential breakdown in the bulk. (b) For +Vg, only electron trapping occurs. The band distortion in the high-*K* bulk would be smaller, leading to higher possibilities of interfacial layer breakdown.

than that caused by the both electrons and holes trapped nearby in the dielectric as shown in Fig. 7.14(a). As a result, the probability of IL- initiated breakdown will be higher if the electric field across the IL is higher. The proposed model which relies on electron and hole trappings within different spatial region in the HfAlO stack is able to successfully explain the polarity dependent Weibull distribution by the energy band distortion inducing preferential breakdown either in the interfacial layer or in the bulk dielectric region.

7.7 Summary

The charge trapping and the breakdown mechanism of HfAlO gate dielectric with TaN metal electrode are investigated. Using carrier separation measurement technique, it is possible to separately monitor the tunneling current components through either bulk high-*K* or interfacial oxynitride layer in an HfAlO/oxynitride stack and determine the breakdown mechanism. It is observed that under negative bias CVS, breakdown is initiated from bulk high-*K* film while positive bias stress tends to initiate interfacial layer breakdown. Statistical distribution of breakdown voltage V_{BD} and charge-to-breakdown Q_{BD} shows two distinct breakdown mechanisms for both negative and positive CVS. Q_{BD} Weibull slope, β for negative bias CVS ranges from 1.5 to 2.1 for p-channel MOS capacitors while β for positive bias CVS is around 0.5 ~ 0.8 for both n-channel and p-channel MOS capacitors. The Weibull slopes obtained for negative CVS is comparable with the β values for the corresponding thickness of HfO_2 [7.37] and Al_2O_3 [7.17] obtained by other researchers while that of positive CVS is very low, showing characteristics of an ultra-thin interfacial layer. The statistical results correlate with the carrier separation measurement results, re-affirming the dual layer breakdown mechanism. The proposed model using charge trappings at different spatial sites within the HfAlO stack could successfully explain the preferential breakdown either in the bulk high-*K* or interfacial layer.

Reference

[7.1] International Technology Roadmap for Semiconductors. (2003). Semiconductor Industry Assoc., San Jose, CA. Available : http://public.itrs.net/Files/2003ITRS/Home2003.htm

[7.2] T. Kauerauf, R. Degraeve, E. Cartier, C. Soens, and G. Groeseneken, "Low Weibull Slope of Breakdown Distributions in High-*K* layers," *IEEE Electron Dev. Lett.*, vol. 23, no. 4, pp. 215-217, 2002.

[7.3] J. W. McPherson, J. Kim, A. Shanware, H. Mogul, and J. Rodriguez, "Trends in the Ultimate Breakdown Strength of High Dielectric-Constant Materials," *IEEE Trans. Electron Dev.,* vol. 50, no. 8, pp. 1771- 1778, 2003.

[7.4] T. Yamaguchi, T. Ino, H. Satake, and N. Fukushima, "Novel dielectric breakdown model of Hf-silicate with high temperature annealing," *Proc. Int. Reliab. Phys. Symp.,* pp. 34-40, 2003.

[7.5] T. Kauerauf, R. Degraeve, E. Cartier, B. Govoreanu, P. Blomme, B. Kaczer, L. Pantisano, A Kerber, and G. Groeseneken, "Towards understanding degradation and breakdown of SiO_2/high-*K* stacks," *IEDM Tech. Dig.,* pp. 521-524, 2002.

[7.6] A. Kerber, E. Cartier, G. Groeseneken, H. E. Maes, and U. Schwalke, "Stress induced charge trapping effects in SiO_2/Al_2O_3 stack with TiN electrodes," *J. Appl. Phys.,* vol. 94, no. 10, pp. 6627-6630, 2003.

[7.7] Z. Xu, M. Houssa, R. Carter, M. Naili, S. D. Gendt, and M. Heyns, "Constant voltage stress induced degradation in HfO_2/SiO_2 gate dielectric stacks," *J. Appl. Phys*., vol. 91, no 12, pp. 10127-10129, 2002.

[7.8] A. Y. Kang and P. M. Lenahan, "Reliability concerns for HfO_2/Si devices : Interface and Dielectric Traps," *Int. Reliab. Workshop*, pp. 102-107, 2002.

[7.9] P. Blomme, B. Govoreanu, J. V. Houdt, and K. D. Meyer, "Temperature dependent current and charge trapping in thick SiO_2/ZrO_2 stacks," *Proc. Int. Reliab. Phys. Symp.,* pp. 208-212, 2003.

[7.10] E. P. Gusev, D. A. Buchanan, E. Cartier, A. Kumar et al., "Ultrathin High-*K* gate stacks for advanced CMOS devices," *IEDM Tech. Dig.,* pp. 451-454, 2001.

[7.11] Z. Xu, M. Houssa, S. D. Gendt, and M. Heyns, "Polarity effect on the temperature dependence of leakage current through HfO_2/SiO_2 gate dielectrics stacks," *Appl. Phys. Lett*., vol. 80, no 11, pp. 1975-1977, 2002.

[7.12] A. Kerber, E. Cartier, R. Degraeve, P. J. Roussel, L. Pantisano, T. Kauerauf, G. Groeseneken, H. E. Maes, and U. Schwalke, "Charge Trapping and Dielectric Reliability of SiO_2-Al_2O_3 Gate Stack with TiN Electrodes," *IEEE Trans. Electron Dev.,* vol. 50, pp. 1261 - 1269, 2003

[7.13] A. Shanware, M. R. Visokay, J. J. Chambers, A. L. P. Rotondaro, H. Bu, M. J. Bevan, R. Khamankar, S. Aur, P. E. Nicollian, J. McPherson, and L. Colombo, "Evaluation of the Positive Biased Temperature Stress Stability in HfSiON Gate Dielectrics," *Proc. Int. Reliab. Phys. Symp.,* pp. 208-212, 2003.

[7.14] S. Zafar, A. Callegari, E. Gusev, and M. V. Fischetti, "Charge Trapping in High K Gate Dielectric Stacks," *IEDM Tech. Dig.,* pp. 217-220, 2002.

[7.15] W. J. Zhu, T. P. Ma, S. Zafar, and T. Tamagawa, "Charge trapping in ultra-thin hafnium oxide," *IEEE Electron Dev. Lett*., vol. 23, pp 597-599, 2002.

[7.16] A. Kerber, E. Cartier, L. Pantisano, R. Degraeve, T. Kauerauf, Y. Kim, A. Hou, G. Groeseneken, H. E. Maes, and U. Schwalke, "Origin of the Threshold Voltage Instability in SiO_2/HfO_2 Dual Layer Gate Dielectrics," *IEEE. Electron Device Lett.,* vol. 24, pp 87 - 89, 2003.

[7.17] A. Kerber, E. Cartier, R. Degraeve, L. Pantisano, Ph. Roussel, and G. Groeseneken, "Strong Correlation between Dielectric Reliability and Charge Trapping in SiO_2/Al_2O_3 Gate Stacks with TiN Electrode," *Proc. Symp. VLSI Technology,* pp 76-77, 2002.

[7.18] R. Degraeve, B. Kaczer, M. Houssa, G. Groeseneken, M. Heyns, "Analysis of high voltage TDDB measurements on Ta_2O_5/SiO_2 stack," *IEDM Tech. Dig.,* pp 327-330, 1999.

[7.19] R. Degraeve, T. Kauerauf, A. Kerber, E. Cartier, B. Govoreanu, Ph. Roussel, L. Pantisano, P. Blomme, B. Kaczer, and G. Groeseneken, "Stress Polarity Dependence of Degradation and breakdown of SiO_2/High-*K* Stacks," *Proc. Int. Reliab. Phys. Symp*., pp. 23-28, 2003.

[7.20] R. Degraeve, A. Kerber, Ph. Roussel, E. Cartier, T. Kauerauf, L. Pantisano, and G. Groeseneken, "Effect of bulk trap density on HfO_2 reliability and yield," *IEDM Tech. Dig.,* pp. 935-938, 2003.

[7.21] W. Zhu and T.P. Ma, "HfO_2 and HfAlO for CMOS: Thermal Stability and Current Transport," *IEDM Tech. Dig*., p. 463, 2001.

[7.22] M. Gutowski, J. E. Jaffe, C. L. Liu, M. Stoker, R. I. Hegde, R. S. Rai, and P. J Tobin, "Thermodynamic stability of high-*K* dielectric metal oxides ZrO_2 and HfO_2 in contact with Si and SiO_2," *Appl. Phys. Lett.,* vol. 80, pp. 1897- 1899, 2002.

[7.23] H. Y. Yu, N. Wu, M. F. Li, C. Zhu, B. J. Cho, D. L. Kwong, C. H. Tung, J. S. Pan, J. W. Chai, W. D. Wang, D. Z. Chi, C. H. Ang, and J. Z. Zheng, "Thermal Stability of $(HfO_2)_x(Al_2O_3)_{1-x}$ on Si," *Appl. Phys. Lett.,* vol. 81, no. 19, pp. 3618-3620, 2002..

[7.24] S. J. Lee, S. J. Rhee, R. Clark, and D. L. Kwong, "Reliability Projection and Polarity Dependence of TDDB for Ultra Thin CVD HfO_2 Gate Dielectrics," *Proc. Symp. VLSI Technology,* pp. 78-79, 2002.

[7.25] B. H. Lee, R. Choi, L. Kang, S. Gopalan, R. Nieh, K. Onishi, Y. Jeon, W. J. Qi, C. S. Kang, and J. C. Lee, "Characteristics of TaN gate MOSFET with ultrathin hafinium oxides (8Å-12Å)," *IEDM Tech. Dig.,* pp. 39-42, 2000.

[7.26] C. H. Lee, J. J. Lee, W. P. Bai, S. H. Bae, J. H. Sim, X. Lei, R. D. Clark, Y. Harada, M. Niwa, and D. L. Kwong, "Self-aligned ultra thin HfO_2 CMOS transistors with high quality CVD TaN gate electrode," *Proc. Symp. VLSI Technology,* pp. 82-83, 2002.

[7.27] R. Choi, C. S. Kang, B. H. Lee, K. Onishi, R. Nieh, S. Gopalan, E. Dharmarajan, and J. C. Lee, " High-Quality Ultra-thin HfO_2 Gate Dielectrics MOSFETs with TaN Electrode and Nitridation Surface Preparation," *Proc. Symp. VLSI Technology*, pp.15-16, 2001.

[7.28] M. S. Joo, B. J. Cho, C. C. Yeo, D. S. H. Chan, S. J. Whoang, S. Mathew, L. K. Bera, N. Bala, and D.-L. Kwong, "Formation of Hafnium-Aluminum-Oxide gate dielectric using single cocktail liquid source in MOCVD process," *IEEE Trans. Electron Dev*., vol. 50, pp. 2088-2094, 2003.

[7.29] K. Yang, Y-C. King, and C. Hu, "Quantum Effect in Oxide Thickness Determination From Capacitance Measurement," *Proc. Symp. VLSI Technology*, Kyoto, Japan, pp. 77-78, June 14-16, 1999.

[7.30] A. Shanware, J. P. Shiely, and H. Z. Massoud, " Extraction of Gate Oxide thickness of n- and p-channel MOSFETs below 20Å from the substrate current resulting from valence-band electron tunneling," *IEDM Tech. Dig.,* pp. 815-818, 1999.

[7.31] S. I. Takagi, N. Yasuda, and A. Toriumi, "Experimental Evidence of inelastic Tunneling and new I-V model for stress-induced leakage current," *IEDM Tech. Dig.,* pp. 323-326, 1996.

[7.32] C. S. Kang, H-J. Cho, Y. H. Kim, R. Choi, K. Onishi, A. Shahriar, and J. C. Lee, "Characterization of resistivity and work function of sputtered-TaN film for gate electrode applications," *J. Vac. Sci. Technol.,* vol. 21, p. 2026, 2003.

[7.33] Y. T. Hou, M. F. Li, H. Y. Yu, Y. Jin, and D.-L. Kwong, "Quantum Tunneling and Scalability of HfO_2 and HfAlO Gate Stacks," *IEDM Tech. Dig.,* pp. 731-734, 2002.

[7.34] M. Houssa, M. Naili, V.V. Afanas, M. M. Heyns, and A. Stesmans, "Electrical and physical characterization of High-*K* dielectrics layers," *Proc. Symp. VLSI Technology.,* pp. 196-199, 2001.

[7.35] M. V. Fischetti, "Generation of positive charge in silicon dioxide during avalanche and tunnel electron injection," *J. Appl. Phys*., vol. 57, no. 8, pp. 2860- 2879, 1985.

[7.36] T. Yamada, M. Moriwaki, Y. Harada, S. Fujii, and K. Eriguchi, "The metal gate MOS reliability with improved sputtering process for gate electrode," *IEDM Tech. Dig*., pp. 319-322, 1999.

[7.37] Y.H. Kim, K. Onishi, C. S. Kang, R. Choi, H.-J. Cho, R. Nieh, J. Han, S. Krishnan, A. Shahriar, and J.C. Lee, "Hard and Soft-Breakdown Characteristics of Ultra-thin HfO_2 Under Dynamic and Constant Voltage Stress," *IEDM Tech. Dig.,* pp. 629-632, 2002.

Chapter 8

Conclusion and Recommendations

8.1 Conclusions

Detailed investigations into the reliability of gate oxide and high-*K* dielectrics have been performed on a wide range of MOS capacitors and transistors. Due to frantic device scaling, gate oxide scaling has been very aggressive in the past few years. Correspondingly, the reliability of oxides in the relevant thickness regime ranging from 45 Å to 13 Å is an ongoing concern and is selected in this thesis for investigation. It was found that the underlying mechanisms of oxide degradation and breakdown change as oxide thickness is scaled downward.

For thin oxide in the thickness regime of 30 Å to 45 Å, the characteristics of gate leakage current prior to quasi-breakdown (pre-QB) and post-QB are investigated. It is observed that hole trapping mechanism in thin oxide (45 Å) is strongly correlated to the QB leakage current. Using bipolar constant current stressing, it is further observed that QB can be characterized into two distinct stages – recoverable and unrecoverable QB. The two stages can be distinguished in terms of electrical recoverability and distinct differences in the V_g-t characteristics. During the recoverable QB stage, two different conduction phases are observed by their disparate carrier separation measurement results. In the first stage, the F-N tunneling electron current is dominant with a small portion of hole current. Subsequent electrical stressing results in the dominance of hole direct tunneling. Using thermal annealing and low stressing current, our studies have also shown that two components of QB leakage current exist: an initial gate leakage at $QB_{threshold}$ and an additional stress-induced component due to post-QB electrical stressing. In the recoverable QB stage, the hole current reduces significantly by a low temperature annealing. The electron current, on the other hand, reduces marginally till its $QB_{threshold}$ level. This implies that the initial leakage component at $QB_{threshold}$ is due to the electron conduction, while the

additional stress-induced component is due mainly to hole conduction. A unifying localized trap region (LTR) model, which is based on energy band distortion due to hole trapping phenomenon near the anode, is proposed. The proposed model is able to explain the current new and past experimental findings. As the conduction model of the unrecoverable QB stage resembles that of the percolation path model, it is believed that most of the controversy may have arisen due to the fact that observations may have been made at different stages of QB.

As gate oxide is degraded electrically to the direct tunneling regime, it was observed that gate leakage in ultra-thin oxides (20 Å) is characterized by step-like increases rather than gradual increases which can be attributed to occurrences of multiple QB. Through direct-current current-voltage (DCIV) measurement, a very good correlation between gate leakage and interface trap irrespective of stressing polarity was observed. It was also observed that quasi-breakdown in thin oxide (20 Å) is distinctly different from that of thicker oxides (45 Å) as shown by the evolution of interface traps and oxide bulk traps at onset of QB. The results suggest that for QB in thin oxide (20 Å), a linked conduction path or percolation path is formed while in thicker oxide, damage is confined to either the anode or cathode without forming a linked path. Using a new criterion based on the gate leakage current increase, it has been observed that lifetime or operating voltage projection becomes significantly worse as device channel area is reduced. Using a simple multiple localized spots model, it is demonstrated that the increased gate leakage current in ultra-thin silicon dioxide can be explained by multiple degraded local spots which are described by Weibull's statistics. It is also estimated that the operating voltage with projected 10-year lifetime, V_{10Y} is about 2.03 V for a 100 μm^2 gate area and 1.9 V for a 10 μm^2 channel area.

Due to high gate leakage, conventional SiO_2 will have to be replaced by high-*K* dielectrics for the 65 nm technology node. Various reliability aspects of high-*K* dielectrics are investigated. It was found that the charge trapping and the breakdown mechanism of HfAlO gate dielectric with TaN metal electrode are polarity dependent. Using the carrier separation measurement technique, which is typically used for electron and hole analysis, it is found that the leakage path through both the bulk and

interfacial layer (IL) of the high-*K* stack can be monitored. This is due to the higher voltage drop across the IL which results in very disparate conduction path for both electrons and holes (or valence electrons for positive gate bias) under negative gate bias. Using carrier separation in this novel method, it is shown that the breakdown mechanism at either bulk high-*K* or interfacial oxynitride layer in an HfAlO/oxynitride stack can be accurately determined. It is further observed that under negative bias CVS, breakdown is initiated from the bulk high-*K* film, while positive bias stress tends to initiate interfacial layer breakdown. Statistical distribution of breakdown voltage V_{BD} and charge-to-breakdown Q_{BD} shows two distinct breakdown mechanisms for both negative and positive CVS. Q_{BD} Weibull slope, β for negative bias CVS ranges from 1.5 to 2.1 while β for positive bias CVS is around 0.5 ~ 0.8. The Weibull slopes for the different gate polarity are similar to those obtained by other researchers and correlate with the carrier separation results. A charge induced breakdown model is proposed and this is correlated to both the charge trapping results and breakdown analysis obtained using the carrier separation method. In this model, it is proposed that hole and electron trapping at different spatial location under negative gate bias resulting in higher electric field in the bulk, thus leading to preferential breakdown in the bulk layer. Conversely, only electron trapping occurs under substrate injection (positive gate bias) and due to the inherently high electric field in the IL, breakdown tends to occur at the IL. The breakdown at different layers of high-*K* stacks show that high-*K* reliability is not entirely similar to SiO_2 due to its material difference and the intrinsic existence of an interfacial layer.

8.2 Recommendations for Future Work

In the course of this study on oxide and high-*K* dielectric reliability, some possibilities were identified for possible future researches and investigations.

In Chapter 4, section 4.3, short bipolar pulses were applied to oxide till quasi-breakdown. The frequency of the pulses applied is relatively small ~ 0.04 Hz. At this low frequency, electrical recovery is observed at onset of QB while subsequent stressing till unrecoverable QB results in a very stable and degraded I-V characteristics. It will be interesting to perform similar bipolar stressing, but using

higher frequency. A frequency dependency study would allows us to selectively probe and de-trap the charges within certain tunneling distance of the degraded oxide, thus giving us a better understanding of the spatial distribution of the traps in the oxide at onset of QB.

In addition, while this thesis has covered most aspects of quasi-breakdown mechanism studies using carrier separation, thermal and bias annealing and carrier modeling after QB, there is still a certain degree of uncertainty due to the complex nature of the breakdown phenomenon. Most researchers have covered the electrical characterization of QB and investigation of its mechanism using an electrical approach. Recently, Pey *et al.* have managed to study the mechanism of complete breakdown using transmission electron microscopy (TEM) [8.1]. Using such physical analysis, very convincing evidence regarding breakdown mechanism can be obtained. While there are attempts to duplicate this for QB, it is unfortunately not successful due to the highly difficult task of proper sample preparation. It is proposed that for future work, TEM on post-QB should be further attempted, to establish clear and convincing evidence for verification and confirmation of QB mechanism.

For high-*K* reliability study in Chapter 7, a novel method to determine IL and bulk breakdown is proposed and verified experimentally. It is observed that our results are not consistent with some of other researchers' result due mainly to difference in the type of dielectrics (Al_2O_3 and ZrO_2), pre-gate treatment and the physical thickness of the bulk and interface layer. For surface preparation and post-deposition treatment, Yang *et al*. have shown that surface nitridation (SN) with NH_3 prior to high-*K* deposition lead to lower EOT, higher leakage current and significantly improved reliability [8.2]. In their CVS stress test, the surface nitrided split shows immunity against quasi-breakdown and hard breakdown (HBD), which they attributed to improved interface quality. The addition of surface nitridation prior to dielectric deposition also results in a thinner interfacial $HfSi_xO_y$ due to the lower diffusion rate of Si through SiN_x which restricts the amount of Si that could diffuse into the Hf overlayer [8.3]. As a result, both bottom and top nitridation results in significantly different material composition which is also the topic of study here. It will be most meaningful if a comprehensive study could be undertaken to study the reliability of

high stacks as its physical thickness scales downward and for various thickness and types of bulk dielectric and interfacial layers.

In recent years, dynamic stressing has become of significant interest due to its close similarity with actual device operation. In normal operation of a p-MOSFET within the framework of a CMOS inverter, the applied gate bias (input signal) is switching between "high" and "low" voltages, while the drain bias (output signal) is exactly out of phase and is alternating between "low" and "high" voltages. Chen *et al*. have performed very detailed studies in the dynamic bias temperature instability (DBTI) of ultra-thin oxides and have found recovery of oxide degradation during the reverse or off pulse [8.4]. Using a pulsed bias stress, it was found that conventional NBTI investigations that were based on static experimental data, have overestimated the degradation of the p-MOSFETs by overlooking the electric passivation (EP) effect during normal operations of the circuits. Similar lifetime improvement using unipolar stressing was also observed in HfO_2 [8.5], although the exact mechanism is still unclear. It would be interesting if similar dynamic, bipolar stressing and positive bias temperature instability (PBTI) technique could be applied to high-*K* to check if similar electrical recovery occurs.

Reference

[8.1] K. L. Pey, C. H. Tung, M. K. Radhakrishnan, L. J. Tang, and W. H. Lin, "Dielectric breakdown induced epitaxy in ultrathin gate oxide - a reliability concern," *IEDM Tech. Dig.,* pp. 163 – 166, 2002.

[8.2] C. W. Yang, Y. K. Fang, S. F. Chen, C. Y. Lin, M. F. Wang, Y. M. Lin, T. H. Hou, L. G. Yao, S. C. Chen, and M. S. Laing, "Effective improvement of high-*K* Hf-silicate/silicon interface with thermal nitridation," *Electronic Lett.,* vol. 39, pp. 421-422, 2003.

[8.3] P. D. Kirsch, C. S. Kang, J. Lozano, J. C. Lee, and J. G. Ekerdt, "Electrical and spectroscopic comparison of HfO_2/Si interface on nitrided and un-nitrided Si(100)," *J. Appl. Phys*., vol. 91, pp. 4345-4363, 2002.

[8.4] G. Chen, M. F. Li, C. H. Ang, J. Z. Zheng, and D. L. Kwong, "Dynamic NBTI of p-MOS transistors and its impact on MOSFET scaling," *IEEE Electron Dev. Lett.,* vol. 23, pp. 734-736, 2002.

[8.5] Y.H. Kim, K. Onishi, C. S. Kang, R. Choi, H.-J. Cho, R. Nieh, J. Han, S. Krishnan, A. Shahriar, and J.C. Lee, "Hard and Soft-Breakdown Characteristics of Ultra-thin HfO_2 Under Dynamic and Constant Voltage Stress," *IEDM Tech. Dig.,* pp. 629-632, 2002.

Appendix A

List of Publications

A.1 Published Journal Papers

1. W. Y. Loh, B. J. Cho, M. S. Joo, M. F. Li, Shajan Matthew, Daniel S.H. Chan, and D.L. Kwong, "Charge Trapping and Breakdown Mechanism in HfAlO/TaN Gate Stack analyzed using Carrier Separation*," accepted for publication in IEEE Dev. Material Reliab., 2004.*

2. W. Y. Loh, B. J. Cho, M. F. Li, Daniel S. H. Chan, C. H. Ang, J. Z. Zhen, and D. L. Kwong, "Localized Oxide Degradatioin in Ultra-Thin Gate Dielectric and its Statistical Analysis," *IEEE Tran. Electron Dev.,* vol. 50, pp. 967 – 972, 2003.

3. C. M. Lek, B. J. Cho, C. H. Ang, S. S. Tan, W. Y. Loh, J. Z. Zhen, and C. Lap, "Impact of decoupled plasma nitridation of ultra-thin gate oxide on the performance of p-channel MOSFETs," *J. Semicond. Sci. Technol*., vol. 17, pp. L25 – L28, 2002.

4. W. Y. Loh, B. J. Cho, and M. F. Li, " Correlation between interface traps and gate oxie leakage current in the direct tunneling regime," *Applied Physics Letters*, vol. 81, no. 2, pp. 379 – 381, 2002.

5. W. Y. Loh, B. J. Cho, and M. F. Li, "Investigation of Quasi-breakdown Mechanism through Post-QB Thermal Annealing," *Japanese Journal of Applied Physics*, vol. 41, pp. 2873 – 2877, 2002.

6. W. Y. Loh, B. J. Cho, and M. F. Li, "Evolution of Quasi-breakdown in Thin Gate Oxides," *Journal of Applied Physics*, vol. 91, no. 8, pp. 5302 – 5306, 2002.

A.2 Conference Papers

7. W. Y. Loh, B. J. Cho, M. S. Joo, M. F. Li, Shajan Matthew, Daniel S. H. Chan, and D. L. Kwong, "Analysis of Charge Trapping and Breakdown Mechanism in High-K Dielectrics with Metal Gate Electrode using Carrier Separation," *IEDM Tech. Dig.,* pp. 927-930, 2003.

8. C. M. Lek, B. J. Cho, W. Y. Loh, C. H. Ang, W. H. Lin, Y. L. Tan, J. Z. Zhen, L. Chan, S. S. Tan, and T. P. Chen, "Effects of Post-Decoupled-Plasma-Nitridation Annealing of Ultra-thin Gate Oxide," *Proc. of 9th IPFA 2002*, pp. 232 – 236.

9. M. F. Li, B. J. Cho, G. Chen, W. Y. Loh, and D. L. Kwong, "New reliability issues of CMOS transistors with 1.3 nm thick gate oxide," *Electrochemical Society Meeting*, Paris, France, April 2003. (Invited)

10. W. Y. Loh, B. J. Cho, M. F. Li, Daniel S. H. Chan, C. H. Ang, Z. J. Zhen, and D. L. Kwong, "Progressive breakdown statistics in ultra-thin silicon dioxides," *Proceeding of 10th IPFA 2003 (IPFA)*, Singapore, 2003.

11. W. Y. Loh, B. J. Cho, M. F. Li, C. M. Lek, Y. F. Yong, and M. S. Joo, "Correlation between Interface Trap and Gate Leakage Current in Ultra-thin Silicon Dioxides," *Proceeding of 9th IPFA 2002*, pp. 246 – 249.

12. W. Y. Loh, B. J. Cho, and M. F. Li, "Bipolar Current Stressing and Electrical Recovery of Quasi-breakdown in thin Gate Oxides," *Proceeding of 8th IPFA 2001*, pp. 59-62.